

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 21:18:11 ; Search time 2502.29 Seconds  
(without alignments)  
8935.871 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgaacatgtccaaaca.....gtagttataaaaaaaa 920

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| C 1        | 120   | 13.0        | 895    | CNS0071A | AL066286 Drosophil |
| C 2        | 113   | 12.3        | 712    | 13       | AX416727 BX416727  |
| C 3        | 111.4 | 12.1        | 1180   | 13       | AX436369 BX436369  |
| C 4        | 104.6 | 11.4        | 946    | 29       | AG127412 Pan trogl |

|      |       |      |      |    |          |
|------|-------|------|------|----|----------|
| C 5  | 103.2 | 11.2 | 922  | 29 | CNS0073W |
| C 6  | 101.8 | 11.1 | 919  | 29 | CNS04ENY |
| C 7  | 100   | 10.9 | 1201 | 9  | AL581589 |
| C 8  | 99.8  | 10.8 | 1201 | 9  | AL531977 |
| C 9  | 99.4  | 10.8 | 1141 | 28 | AQ743305 |
| C 10 | 99    | 10.8 | 1101 | 29 | CNS00FXE |
| C 11 | 98    | 10.7 | 1151 | 14 | CD501049 |
| C 12 | 97.2  | 10.6 | 1200 | 9  | AL548181 |
| C 13 | 96.6  | 10.5 | 1024 | 29 | AG133080 |
| C 14 | 96.2  | 10.5 | 753  | 13 | EX433342 |
| C 15 | 96.2  | 10.5 | 1100 | 13 | BX385076 |
| C 16 | 96.2  | 10.5 | 1172 | 29 | BZ696157 |
| C 17 | 95.8  | 10.4 | 1097 | 29 | CNS01224 |
| C 18 | 95.8  | 10.4 | 1101 | 29 | CNS0100X |
| C 19 | 95    | 10.3 | 791  | 29 | CNS009KS |
| C 20 | 95    | 10.3 | 1101 | 29 | CNS00LQO |
| C 21 | 95    | 10.3 | 1144 | 28 | AQ743364 |
| C 22 | 94.6  | 10.3 | 899  | 13 | AX436853 |
| C 23 | 94.4  | 10.3 | 1216 | 29 | AG135357 |
| C 24 | 94.2  | 10.2 | 815  | 28 | B12686   |
| C 25 | 94    | 10.2 | 1000 | 13 | EX418086 |
| C 26 | 94    | 10.2 | 1188 | 29 | AG135332 |
| C 27 | 93.8  | 10.2 | 859  | 29 | AG128925 |
| C 28 | 93.8  | 10.2 | 1008 | 29 | AG137085 |
| C 29 | 93.8  | 10.2 | 1099 | 28 | AQ743360 |
| C 30 | 93.6  | 10.2 | 884  | 29 | CNS008UO |
| C 31 | 93.6  | 10.2 | 969  | 28 | AQ743309 |
| C 32 | 93.6  | 10.2 | 1081 | 29 | AG135328 |
| C 33 | 92.8  | 10.1 | 955  | 29 | AG076494 |
| C 34 | 92.8  | 10.1 | 981  | 29 | AG127518 |
| C 35 | 92.8  | 10.1 | 1183 | 29 | BZ696114 |
| C 36 | 92.8  | 10.1 | 1353 | 29 | AG128010 |
| C 37 | 92.6  | 10.1 | 923  | 13 | EX410248 |
| C 38 | 92.6  | 10.1 | 1183 | 29 | AG136828 |
| C 39 | 92.4  | 10.0 | 880  | 29 | AG139490 |
| C 40 | 92.2  | 10.0 | 885  | 29 | CNS031TM |
| C 41 | 92    | 10.0 | 1101 | 29 | CNS00LT2 |
| C 42 | 92    | 10.0 | 1137 | 28 | AQ743326 |
| C 43 | 92    | 10.0 | 1225 | 29 | CNS0166K |
| C 44 | 92    | 10.0 | 1491 | 10 | BE882936 |
| C 45 | 91.4  | 9.9  | 700  | 29 | AG127423 |

#### ALIGNMENTS

RESULT 1  
CNS0071A/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BAC14B09 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION  
AL066286  
VERSION  
GSS  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE  
1 (bases 1 to 895)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammos in Pieter de Jong's laboratory in the Department of



[illegible]

```

QY 520 ACCCTGAGCATATGAGCAACAAATCCAACTGCTGGCTATCTGCAAAAGAAATACCAA 579
Db 453 AAAACAACCAATCTCTCCCAAAACAAACAAACAAACAAACAAACAAACAAACAAAC 512
QY 580 CAAAAAACAGGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 639
Db 513 AAAAAAACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 572
QY 640 CAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACT 699
Db 573 GAAAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 632
QY 700 AGAAGAGCCCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 759
Db 633 AAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 692
QY 760 CACCACAGGAATCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAA 819
Db 693 AAACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 752
QY 820 AGGCAATCTAAGCCCTTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAG 879
Db 753 CAGCACCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 812
QY 880 TCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 919
Db 813 CCAAAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 852

```

RESULT 5  
CNS0073W/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster (fruit fly)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 922)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osoegawa and  
Aaron Mammeter in Peter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
source  
1..922  
/organism="Drosophila melanogaster"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR14D09"  
/note="end : TET3"

BASE COUNT  
223 a 95 c 109 g 221 t 274 others

## ORIGIN

Query Match 11.2%; Score 103.2; DB 29; Length 922;  
Best Local Similarity 24.8%; Pred. No. 2.3e-07;  
Matches 102; Conservative 164; Mismatches 144; Indels 1; Gaps 1;  
QY 386 CAACCCACACACTCAAGACTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 445  
Db 921 MMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMW 862  
QY 446 ACTTACAAACAAACGCGCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 505  
Db 861 ACANWAAWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMW 802  
QY 506 GTGTTTAACTTTGTACCTGTCAGCATATGTCAGCAACAAATCCAACTGCTGGCTATCTGC 565  
Db 801 ACACWCAWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMW 743  
QY 566 AAAAGATATACCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 625  
Db 742 AMAAMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 683  
QY 626 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 685  
Db 682 MWCAWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 623  
QY 686 ACCACCAAGCCCAACAAAGAGCCCAACCAATCAACCAACCAACCAACCAACCAACCAAC 745  
Db 622 AMWACMAWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMW 563  
QY 746 CTGCTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 796  
Db 562 AMWACMAWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMW 512

RESULT 6  
CNS04ENY/cLOCUS  
DEFINITION

919 bp DNA linear GSS 01-SEP-2000  
Tetraodon nigroviridis genome survey sequence T7 end of clone  
104P14 of library G from Tetraodon nigroviridis, genomic survey  
sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae; Tetraodon.

## REFERENCE

## AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouteau, L., Fisher, C.,  
Bernot, A., Fzames, C., Wincker, P., Brottier, P., Quetier, F.,  
Saurin, W. and Weissenbach, J.

## TITLE

Estimate of human gene number provided by genome-wide analysis  
using tetraodon nigroviridis DNA sequence

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,  
Fizames, C., Fischer, C., Bouteau, L., Billault, A., Quetier, F.,  
Saurin, W., Bernot, A. and Weissenbach, J.

## TITLE

Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,  
Fizames, C., Fischer, C., Bouteau, L., Billault, A., Quetier, F.,  
Saurin, W., Bernot, A. and Weissenbach, J.

## TITLE

Genoscope - Centre National de Sequencage :  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## Direct Submission



was not normalized. Library was constructed by Life Technologies division of Invitrogen. This sequence belongs to sequence cluster 8262.f. For more information about this cluster, see <http://www.genoscope.cns.fr/>.  
cgt-bio/cluster.cgi?seq=CS0DG004CC08NP1&cluster=8262.f. Contact : Bing Liang Email : [fliang@life.techn.com](mailto:fliang@life.techn.com) URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DG004CC08NP1

rarasay Avenue Genoscope sequence ID : CS0DG004CC08NP1.  
 Location/Qualifiers  
 1..1201

| Location/Qualifiers |
|---------------------|
| 1. 1201             |

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG004YF15"
/tissue_type="B CELLS"
(DAMOS CBUT LINE)
```

92 a 57 c 490 g 461 t 101 others

th 10.9%; Score 100; DB 9; Length 1201;  
Similarity 45.5%; Pred. No. 7,2e-07;  
56; Conservative 39; Mismatches 264; Indels 4; Gaps 2;  
7 AACACACCACTACTAGCTTCAACAACACCGAGCTCAAGTCAAACTGCAACCCCAAC 396

2 AAAAAAAAAA CAAAAANAAAAAA CAAGNAAACCA NNAAAAAAAAAAAAAAAAAAAAAA CAAAAA 1093  
 7 AGTCAGA GCTAAAAACACACAACCAACCCCAACCAAGCCGCACTTACAAAACA 456  
 2 AACCNVCAAAAAAANAACAAAAAANAANAACAAAAACAAAAAACAAAAAATAAC 1033  
 7 ACGCCAAAAACAACCACCAAACCAACCCCAATGAATTTTCACTTCGAAGTGCTTAACATT 516

7 TGTACCTCGAGCATATGCGACCAATCCAACTGCTGGGGTATCTGCAAAAAGAATACC 576  
2 AAACACAAACCCCAAAATAAACACCAACCAAAACAACCAACCAACCAACCAACCAACCAAC 913

7 AATCAAAAAGTCTCAAGCCTTAAACCACTTCAAGAC 636  
 1 AATCAAAAAGTCTCAAGCCTTAAACCACTTCAAGAC 636  
 2 MWCCMACAACAMMAAAAAAAAAACCA--AACAAACCAAAAAAAAAACCMWCMWMAWA 855  
 7 AACCAAAAAGTCTCAAGCCTTAAACCACTTAAACCAAGGAGTACCCACCAACGCC 696  
 1 AAC--AACACACCCCAAAAAAAMWMAAAAAACAAAAACAAACAAACAAACAAACAA 797  
 4 AAC--AACACACCCCAAAAAAAMWMAAAAAACAAAAACAAACAAACAAACAAACAA 797

7 CACAGAGAGCGCAACCATCAACACCAACCAAAACAAACATCAACAACCTACACTGCTCACCA 756  
6 AAAAAACAAATTCACAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 737  
7 CAACACCAACGAGAAATCCAAAACTCACAGCTCAATGGAACCTTCCACTCAACTCCCTC 816

6 MMAACACCCCAACCAAAATAAAACAAMMAAAACACMMAAACAAACCCCAACACCCCAACCM 677

7 CGAAGCAATCTAAGCCCTTCTCAAGTCTCCCAACATCCGAGCACCCGATCAACACCTC 876

8 CACCCGCCCAACACACCCCCCAACACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC 617

9 ATCTCCACCCCAACACCAACGCC 899

6 CCCCCCCCCCCCCCCCCCCC 594





Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.  
1 (bases 1 to 1151)  
Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M., Schmutz, J., and Myers, R.M.  
Expressed sequence tags from *Gasterosteus aculeatus*  
Unpublished  
Contact: Kingsley, DM  
HMI and Department of Developmental Biology  
Stanford University School of Medicine  
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA  
Tel: 650 725 5954  
Fax: 650 725 7739  
Email: kingsley@cmgm.stanford.edu  
Plate: 48  
High quality sequence start: 3  
High quality sequence stop: 263.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

1..1151  
/organism="Gasterosteus aculeatus"  
/mol\_type="mRNA"  
/strain="Galinas river, CA"  
/db\_xref="taxon:69293"  
/clone="CDA48-P07"  
/sex="mixed male and female"  
/tissue\_type="heads and internal organs combined"  
/dev\_stage="adult"  
/clone\_lib="SHGC-CDA"  
/note="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI (5' adaptor); Site 2: XhoI (3' linker primer); The mixed organ cDNA library was generated using the ZAP-CDNA method by Stratagene. First strand cDNA synthesis was primed with a 50 bp linker primer containing an oligo dT sequence preceded by a synthetic XhoI site. 5 prime adaptors were used containing an EcoRI cohesive end. The finished cDNAs were inserted in to the ZAP express vector unidirectionally in the sense orientation with respect to the lacZ promoter of pBK-CMV. An amplified library was prepared from approximately 3 million primary clones in the lambda ZAP Express vector. In vivo excision was then used to generate individual pBK-CMV phagemid clones for EST sequencing."  
EST sequencing." 73 a 51 c 391 g 530 t 106 others

## BASE COUNT

## ORIGIN

Query Match 10.7%; Score 98; DB 14; Length 1151;  
Best Local Similarity 45.7%; Pred.No. 1.5e-06;  
Matches 319; Conservative 0; Mismatches 375; Indels 4; Gaps 1;  
QY 202 AAACCAAGTGCACCTAAGTGCATATCATAGATGCACACAGCCAGATCAAGAA 261  
1084 AAACCAAGTGCACCTAAGTGCATATCATAGATGCACACAGCCAGATCAAGAA 1025  
QY 262 CACAAACCCACATACCTCAGATGCTTGGATGCTTCCATCTGTC 321  
1024 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 965  
QY 322 TGAATATCATC-----ACAAACCAACCAACCAACCAACCAACCAACCAACCAAC 377  
964 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 905  
QY 378 CAAACCTGCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 437  
904 CAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 845  
QY 438 GCAAGCCCACTACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 497  
844 ACCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 785  
QY 498 ACTTCGAAGTGTATTTTGTATCTTGTACCTGTCAGCATATGATGATGATGATG 557

Db 784 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 725  
QY 558 CTATCTGCAAAAGATATCAACCAACCAACCAACCAACCAACCAACCAACCAAC 617  
Db 724 AAACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 665  
QY 618 AAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677  
Db 664 CAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 605  
QY 678 AAGTACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 737  
Db 604 CCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 545  
QY 738 CAACCTACCTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 797  
Db 544 CAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 485  
QY 798 CTTTCCACTCAACCTCTCGAAGGCAATCTAGCCCTTCTCAAGTCTTCAACATCC 857  
Db 484 CACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 425  
QY 858 AGCACCACATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 895  
Db 424 CCCCCCCCCNN 387

## RESULT 12

## AL548181

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

On Feb 15, 2001 this sequence version replaced gi:12882945.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 113.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0D1034BC010P1&cluster=113.r. Contact :

Peng Liang Email: fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0D1034BC010P1.

## FEATURES

## source

1..1200

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1034YE02"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

447 a 362 c 185 g 95 t 111 others

## BASE COUNT

## ORIGIN

## Query Match

## Best Local Similarity

## 40.1%;

## Pred. No. 2e-06;

## Length 1200;

```
Matches 252; Conservative 82; Mismatches 279; Indels 15; Gaps 2;
QY 172 AATTACAGCCATCATATTCATAGCTCGGCAACCAACAAAGTCACACTAACTGCAAT 231
Db |||||
QY 573 AAGCAGACGACAACTCTTCTTCCAAAAAATAAATAAATAAATAAATAAATAA 632
Db |||||
QY 232 CATACAGATGCAACAGCAGATCAAGACACAAACCCCAACATACCTCAGGATCC 291
Db |||||
QY 633 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 692
QY 292 TCAGCTTGGAATCAGCTTCTCCATCTGTCTGAAATTCATCAACACCACTACT 351
Db |||||
QY 693 CCWCCCCCHCTTCCCCCCCCCAAHCHCCMACACACACACCAACAAACCA 752
QY 352 AGCTTCAACACACACAGAGTCAAGTCAAACTGCAACCCCAACAGCTCAAGCT 411
Db |||||
QY 753 MAAAAACACACACACACACACACACACACACACACACACACACACACAC 812
QY 412 CACACACACACACACACACACACACACACACACACACACACACACACAC 464
Db |||||
QY 813 CACACACACACACACACACACACACACACACACACACACACACACACAC 872
QY 465 -----ACACACACACACACACACACACACACACACACACACACACAC 516
Db |||||
QY 873 YAAAAACACACACACACACACACACACACACACACACACACACACACAC 932
QY 517 TGTACCTGAGCATATGACGACCAACATCCCACTGCTGGGTATCTGCAAAAGATACC 576
Db |||||
QY 933 ACACACACACACACACACACACACACACACACACACACACACACACAC 992
QY 577 AACAAAAACACAGGAAGAAACACACACACACACACACACACACACACACAC 636
Db |||||
QY 993 AACACACACACACACACACACACACACACACACACACACACACACACAC 1052
QY 637 AACCAAAAGATCTCAACCTCAACACACACACACACACACACACACACACAC 696
Db |||||
QY 1053 MAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1112
QY 697 CACAGAGGACCAACCATCAACACACACACACACACACACACACACACACAC 756
Db |||||
QY 1113 AACAAAAACACAGGAAGAAACACACACACACACACACACACACACACACAC 1172
QY 757 CACACACACAGGAATTCACAACTCACA 784
Db |||||
QY 1173 CACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1200

RESULT 13
AG133080
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AG133080
Pan troglodytes DNA, clone: PTB-145K08.R, genomic survey sequence.
AG133080
AG133080.1 GI:16662758
GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 1024)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanse@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
COMMENT
```

was generated during the R&D process and may have higher chance of  
clone tracking errors.

## PRIMERS

Sequencing: M13Rev

## LIBRARY

Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI  
Location/Qualifiers  
1. .1024

/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-145K08.R"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"  
495 a 460 c 23 g 40 t 6 others

## FEATURES

source

BASE COUNT  
ORIGIN

Query Match  
Best Local Similarity  
Matches 354; Conservative 0; Mismatches 390; Indels 4; Gaps 2;

10.5%; Score 96.6; DB 29; Length 1024;  
47.3%; Pred. No. 2.6e-06;  
0; Mismatches 390; Indels 4; Gaps 2;

QY 155 ATCTCACTTCACTTATTAATTACAGCCATCATATTCATAGCTCGGCAACCACTGCAAAAGTC 214

Db 202 AACAAACCCCAACACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC 281

QY 215 ACACAAACCTGCAATCATACAGATGCAACAGCCAGATCAAGACCAACCAACCAACCAACCAACCA 274

Db 262 AACCCCAACCAACCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC 321

QY 275 TACCTCACTCAGATGCTCAGCTTGGATTCAGTCTTCCAAATCTGTCTGAAATTCATCA 334

Db 322 ACCCCCAACCAACCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC 381

QY 335 CAAACCAACCACTACTAGCTTCAACACACACAGGAGTCAAGTCAACCTGCAACCAACCA 394

Db 382 ACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 441

QY 395 ACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 454

Db 442 CCCCCCAACCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 501

QY 455 CAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 512

Db 502 CAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 561

QY 513 ACTTTGTACCTGAGCATATGACGACCAACATCCCAACCTGCTGGGTATCTGCAAAAGTC 572

Db 562 ACCCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 621

QY 573 TACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 632

Db 622 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 681

QY 633 AGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 692

Db 682 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 741

QY 693 AGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 752

Db 742 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 801

QY 753 CCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 812

Db 802 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 859

QY 813 CTTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 872

Db 860 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 919

QY 873 CTTCACTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 900

```

Db      920 CATCACACACACACACACACCCCA 947
      ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 14
BX443342/c
LOCUS
DEFINITION
BX443342 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
clone CS0DH001YF04 3-PRIME, mRNA sequence.
ACCESSION
BX443342
VERSION
BX443342.1 GI:30782114
SOURCE
EST.
ORGANISM
Homo sapiens (human)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 753)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3806.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DH001DC02NP1&cluster=3806.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DH001DC02NP1.
FEATURES
Location/Qualifiers
source
1..753
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH001YF04"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 53 a 29 c 99 g 264 t 308 others
ORIGIN
Query Match
Best Local Similarity 10.5%; Score 96.2; DB 13; Length 753;
Matches 176; Conservative 122; Mismatches 342; Indels 3; Gaps 1;
QY 234 TACAAGATGCAACAGCCAGATCAAGAACACACACCAACCACTCACTCAGGATCTCTC 293
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 753 TMMGGGMMGVMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMM 694
QY 294 AGCTTGGAATCAGCTTCTCCATCTGCTGAAATATCATCAACCAACCAACCACTAG 353
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 693 MMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMM 634
QY 354 CTTCAACACACACAGGAGTCAAGTCAACCTGCAACCCACCAACAGTCAAGACTAAACA 413
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 633 MMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMM 574
QY 414 CAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 473
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 573 MMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMMGMM 514
QY 474 CAACACACCAACCAATGATTTTCACTTCGAGAGTGTTTAACTTTGTACCTCGAGCATAT 533
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 513 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 454

```

```

QY 534 GCACCAACATCCAACTGCTGGCTATCTGCAAAAGAAATACCAACAAACCAAGAGAA 593
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 453 NCANNNCNCAACGCGCCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNC 394
QY 594 AGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 653
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 393 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 334
QY 654 AACCTCAACCACTAAACCAAGAGAGTACCCACCACCAAGCCCAAGAGAGAGCAACCA 713
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 333 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 274
QY 714 TCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 773
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 273 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 214
QY 774 CAAACTCACAAGTCAAAATGGAACCTTCACACTCAACCTCTCCGAGAGGCAATCTAAGCC 833
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 213 MAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 157
QY 834 CTTCTCAAGTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTC 876
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 114
RESULT 15
BX385076
LOCUS
DEFINITION
BX385076 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL006YE12 3-PRIME, mRNA sequence.
ACCESSION
BX385076
VERSION
BX385076.1 GI:30436528
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1100)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4073.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL006BC06NP1&cluster=4073.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL006BC06NP1.
FEATURES
Location/Qualifiers
source
1..1100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL006YE12"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched double-strand cDNA was
digested with Not I and cloned into the Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 433 a 175 c 126 g 196 t 170 others
ORIGIN
Query Match
Best Local Similarity 10.5%; Score 96.2; DB 13; Length 1100;
Matches 167; Conservative 72; Mismatches 190; Indels 0; Gaps 0;

```



```
Qy 351 TAGCTTCAACCAACACCGAGTCAAGTCAAACTGCAACCCCAACCAACAGTCAAGACTAAAA 410
Db 672 TWWAWWWAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 731
Qy 411 ACACAAACACACACACACACACACACACACACACACACACACACACACACACACACAC 470
Db 732 AAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 791
Qy 471 CACCAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 530
Db 792 AAATAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 851
Qy 531 TATGACGCAACATCCACCTGCTGGGTATCTGCAAAAGAAATACCAAAACAAATAAACAC 590
Db 852 AAAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 911
Qy 591 GAAGAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 650
Db 912 MAMMACAAATAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 971
Qy 651 TCAAACTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 710
Db 972 MAAAMCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1031
Qy 711 CCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 770
Db 1032 AAAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1091
Qy 771 ATCCAAAC 779
Db 1092 AAAMCMCC 1100
```

Search completed: October 30, 2003, 01:15:23  
Job time : 2507.29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2003, 00:01:13 ; Search time 310.043 Seconds  
(without alignments)  
8070.528 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgcacacatgtccaaaca.....gtagttattataaaaaaaa 920

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1811591 seqs, 135986290 residues

Total number of hits satisfying chosen parameters: 3623182

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length  | ID | Description         |
|------------|-------|-------------|---------|----|---------------------|
| 1          | 853.4 | 92.8        | 15223   | 10 | US-09-847-173-1     |
| 2          | 637.4 | 69.3        | 696     | 10 | US-09-844-645-4     |
| 3          | 441.4 | 48.0        | 15225   | 12 | US-09-827-688-10    |
| C 4        | 86.2  | 9.4         | 7758    | 12 | US-10-311-455-1076  |
| C 5        | 85.8  | 9.3         | 793     | 10 | US-09-878-574-4304  |
| C 6        | 85.8  | 9.3         | 3673778 | 12 | US-10-312-841-2     |
| C 7        | 84.2  | 9.2         | 1372    | 12 | US-10-017-161-1981  |
| C 8        | 81.4  | 8.8         | 16167   | 12 | US-10-311-455-1056  |
| C 9        | 81.4  | 8.8         | 16167   | 12 | US-10-240-485-82    |
| C 10       | 81    | 8.8         | 6270    | 12 | US-10-311-455-378   |
| C 11       | 81    | 8.8         | 6270    | 12 | US-10-240-453-20    |
| C 12       | 81    | 8.8         | 6270    | 14 | US-10-239-676-12    |
| C 13       | 80    | 8.7         | 6668    | 12 | US-10-311-455-1670  |
| C 14       | 78.2  | 8.5         | 1635    | 9  | US-09-864-761-20241 |
| C 15       | 78.2  | 8.5         | 1973    | 9  | US-09-864-761-3471  |
| C 16       | 78.2  | 8.5         | 2543    | 12 | US-10-029-386-20536 |

Sequence 4296, Ap  
Sequence 2109, Ap  
Sequence 4306, Ap  
Sequence 19241, A  
Sequence 2513, Ap  
Sequence 1669, Ap  
Sequence 4312, Ap  
Sequence 675, App  
Sequence 648, App  
Sequence 1931, App  
Sequence 4315, Ap  
Sequence 25133, A  
Sequence 20723, A  
Sequence 1241, Ap  
Sequence 2004, Ap  
Sequence 1783, Ap  
Sequence 187, App  
Sequence 165, App  
Sequence 1857, Ap  
Sequence 1, Appli  
Sequence 1793, Ap  
Sequence 45, Appl  
Sequence 401, App  
Sequence 268, App  
Sequence 1123, Ap  
Sequence 54, Appl  
Sequence 52, Appl  
Sequence 327, App  
Sequence 78, Appl

US-09-878-574-4296  
US-09-983-965-2109  
US-09-878-574-4306  
US-09-864-761-19241  
US-09-864-761-19241  
US-10-311-455-1669  
US-09-878-574-4312  
US-10-311-455-675  
US-10-311-455-648  
US-10-311-455-1931  
US-09-878-574-4316  
US-10-029-386-25133  
US-10-029-386-20723  
US-10-311-455-1241  
US-10-311-455-2004  
US-10-311-455-1783  
US-10-240-453-187  
US-10-239-676-165  
US-10-017-161-1857  
US-10-312-841-1  
US-10-311-455-1793  
US-10-240-485-45  
US-10-311-455-401  
US-10-311-455-268  
US-10-311-455-1123  
US-10-240-453-54  
US-10-239-676-52  
US-10-240-453-327  
US-10-311-455-78

#### ALIGNMENTS

#### RESULT 1

US-09-847-173-1

; Sequence 1, Application US/09847173

; Publication No. US20020182228A1

; GENERAL INFORMATION:

APPLICANT: Collins, Peter L.

TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY

SYNCYTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: CA

COUNTRY: U.S.A.

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/847,173

FILING DATE: 03-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/720,132

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 15280-250-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-467-9600

TELEFAX: 415-576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 15223 base pairs

TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-847-173-1

Query Match      92.8%; Score 853.4; DB 10; Length 15223;
Best Local Similarity 95.5%; Pred. No. 1.8e-210;
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCAAAACATGTCACAAACCAAGACCAACGACCGCTAAGACACTAGAAAGACCTGGGA 60
DB 4682 TGCACACATGTCACAAACCAAGACCAACGACCGCTAAGACACTAGAAAGACCTGGGA 4741
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT 120
DB 4742 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT 4801
QY 121 AGCACAAATCACATTTATTCATATTCGCAATGATTAATCTCACTTATATATTACAGC 180
DB 4802 AGCACAAATCACATTTATTCATATTCGCAATGATTAATCTCACTTATATATTGAGC 4861
QY 181 CATCATATTCATAGCTCGGCAACCAACCAAGGTCACAACTCACTGCAATCATACAAGA 240
DB 4862 CATCATATTCATAGCTCGGCAACCAACCAAGGTCACAACTCACTGCAATCATACAAGA 4921
QY 241 TGCACCAAGCGAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300
DB 4922 TGCACCAAGCGAGATCAAGAACCAACCCCAACATACCTCAGGATCTCAGCTTGG 4981
QY 301 AATCAGCTTCTCAATCTGCTGAAATTAATACATACCAACCAACCACTAGCTTCAAC 360
DB 4982 AATCAGCTTCTCAATCTGCTGAAATTAATACATACCAACCAACCACTAGCTTCAAC 5041
QY 361 AACACCAAGGAGTCAAGTCAAACTGCAACCAACCAAGTCAAGTCAAACTCAAAACCAAC 420
DB 5042 AACACCAAGGAGTCAAGTCAAACTGCAACCAACCAAGTCAAGTCAAACTCAAAACCAAC 5101
QY 421 AACCAAAACCAACCCAGAGGCGCATCAAAACCAAGGCGCAAAACCAACCAACCAACAA 480
DB 5102 AACCTCAACCAACCAACCCAGAGGCGCATCAAAACCAAGGCGCAAAACCAACCAACCAAC 5161
QY 481 ACCCAATTAATGTTTTCACCTCGAGTGTTCATTTGTCCTGTCAGCATATGCGACAA 540
DB 5162 ACCCAATTAATGTTTTCACCTCGAGTGTTCATTTGTCCTGTCAGCATATGCGACAA 5221
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGAAAGAAAC 600
DB 5222 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGAAAGAAAC 5281
QY 601 CACCAACCAAGCTCAAAACCAACCACTTCAGACAAACCAAAAGATCTCAACCTCA 660
DB 5282 CACTTACCAAGCCCAACCAAAACCAACCTTCAGACAAACCAAAAGATCTCAACCTCA 5341
QY 661 AACCACTTAAACCAAGGAGTACCCACCAACCAAGCCCAAGAGGCGCAACCAACCAAC 720
DB 5342 AACCACTTAAACCAAGGAGTACCCACCAACCAAGCCCAAGAGGCGCAACCAACCAAC 5401
QY 721 CACCAAAACCAACCAATCACTACCTGCTCAACCAACCAACCAACCAAGAAATCCAAACT 780
DB 5402 CACCAAAACCAACCAATCACTACCTGCTCAACCAACCAACCAACCAAGAAATCCAAACT 5461
QY 781 CAAAGTCAATGAAACCTTTCCCTCACTCACTCTCCGGAAGGCAATCTAAGCCCTTCTCA 840
DB 5462 CAAAGTCAATGAAACCTTTCCCTCACTCACTCTCCGGAAGGCAATCTAAGCCCTTCTCA 5521
QY 841 AGTCTCCACACATCCGAGGACCCATCACACCTCTCTCTCCACCCCAACCAACCAACG 900
DB 5522 AGTCTCTACACATCCGAGGATCCCATCACACCTCTCTCTCTCCACCCCAACCAACCAACG 900
QY 901 GTAGTTATTTAAAAA 919
DB 5582 GTAGTTATTTAAAAACATA 5600
```

```

RESULT 2
US-09-844-645-4
; Sequence 4, Application US/09844645
; Patent No. US20020102242A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; Curiel, David T.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
; ADMINISTERING PNEUMOCOCCAL DNA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,645
; FILING DATE: 04-DEC-1996
; FILING DATE: 27-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,505
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-844-645-4

Query Match      69.3%; Score 637.4; DB 10; Length 696;
Best Local Similarity 96.1%; Pred. No. 3.7e-155;
Matches 684; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 8 ATGTCCAAAACCAAGACCAACGACCGCTAAGACACTAGAAAGACCTGGGACACTCTC 67
DB 1 ATGTCCAAAACCAAGACCAACGACCGCTAAGACACTAGAAAGACCTGGGACACTCTC 60
QY 68 AATCATTTATTTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTTAGCAAA 127
DB 61 AATCATTTATTTATTCATATCATCGGTCTTATATAAGTTAAATCTTAAATCTTAGCAAA 120
QY 128 ATCACATTTATTCATTTCTGGCAATGATAATCTCACTTCTTATATTAATACAGCCATCATA 187
DB 121 ATCACATTTATTCATTTCTGGCAATGATAATCTCACTTCTTATATTAATACAGCCATCATA 180
QY 188 TTCAATAGCTCGGCAACCAACCAAGTCAACCTAACTGCAATCATATAAGATGCAACA 247
DB 181 TTCAATAGCTCGGCAACCAACCAAGTCAACCTAACTGCAATCATATAAGATGCAACA 240
QY 248 AGCCAGATCAAGAACCAACCAACCAACCAACCTCACTCACTGAGATCTCAGTTGGAATCAGC 307
DB 241 AGCCAGATCAAGAACCAACCAACCAACCAACCTCACTCACTGAGATCTCAGTTGGAATCAGT 300
QY 308 TTCTCCAAATCTGCTGAAATTAATCATCAAAACCAACCACTACTAGCTTCAACACCA 367
```



US-10-311-455-1076

Query Match  
Best Local Similarity 9.4%; Score 86.2; DB 12; Length 7758;  
Matches 312; Conservative 0; Mismatches 338; Indels 3; Gaps 2;  
QY 121 AGCACAATTCATATTCATTCGCAATGATAATCTCAACTTCACTTTATATTTACAGC 180  
DB 1866 AACAAAAAACAACCAAACTCCCAAAACAAATTAATAAATAAAGACCTTAAC 1807  
QY 181 CATCATATTCATAGCTCGGCAAAACCAAAAGTCACTTAACAATGCAATCATACAAGA 240  
DB 1806 AACAAAAAACAACCTTAACATCTTAAACCAAAAAAACAACCTTAAACCAACTTAA 1747  
QY 241 TGCACAGCCGATCAAGACACAAACCCCAACATCTCACTAGGATCTCTCAGCTGG 300  
DB 1746 AACAAAAAACAACCTTAACACCTTAAACCAAAAAAACAACCTTAAACCACTTAA 1687  
QY 301 AATCAGCTTCTCAATCTGTCTGAAATTTACATCAAAACCCACCATAGTCTTCAAC 360  
DB 1686 ACCAAAAAACAATTAACCTCCCAAAACAAATTAATTAATTAACCTTAAACCA 1627  
QY 361 AACACGAGGTCATAGTCAAACTCGCAACCCCAACAGTCAAGACTTAAACCAACAC 420  
DB 1626 AACCAACAAAAACAACCAAAACACCTTAAACAAAAAACAACCTTAAACCACTTAA 1567  
QY 421 AACCCAAACACCAACCCAGCCCACTTACAAACCAACGCGCAAAACCAACCAACAA 480  
DB 1566 AACCAAAAAAACAACCTTAACATCTTAATACCAAAAAAACAACCTTAAACCAAC 1507  
QY 481 ACCCAAT--AATGATTTTCACTTCGAAGTGTAACTTTGACCTGCGAGCATATGCG 538  
DB 1506 AACAAATAAACCTTAACCAACAAATTAACAACTTAACACCTTAAACCAAAAA 1447  
QY 539 AACAAATCAACCTGTGGCTATCTGC-AAAAGAAATACCAACAAAAAACCCAGGAAGA 597  
DB 1446 AACAACTAAACATCTAAACCAACCAACAAACAAATTAACAAATTCACAAACAA 1387  
QY 598 AACCAACCAACGCTTACAAAAAACCACTTCAAGACACCAAAAAAGATCTCAAC 657  
DB 1386 AACCAACAAAAACCTTAACCAAAAAAACAACCTTAAACCTTAAACCAACCAAAA 1327  
QY 658 TCAAAACCACTAAACCAAGGAGTACCCACCAACGCGCCACAGAGGCGCAACCATCA 717  
DB 1326 ACCAAATAAACCTTAAACCAAAAAAACAACAACTTAACCTTCCCAAAACCAACAA 1267  
QY 718 CACCAACCAAAACATCACTCACTGCTGCTACCAACCAACCAACCAACCAAGGA 770  
DB 1266 AACCAACCAAAACCTTAAACCAAAAAAACAACCAACCACTTAAACCAACCAAAA 1214

RESULT 5  
US-09-878-574-4304/c  
; Sequence 4304, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 4304  
; LENGTH: 793  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure

; LOCATION: (1)..  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-C11  
US-09-878-574-4304

Query Match  
Best Local Similarity 9.3%; Score 85.8; DB 10; Length 793;  
Matches 283; Conservative 0; Mismatches 300; Indels 2; Gaps 2;  
QY 191 ATAGCTCGGCAAAACCAAAAGTCACTTAACAATGCAATCATACAAGATGCAACAGC 250  
DB 715 ACAAACACACCAACCAACCAAAACCAACCAACCAACCAACCAACCAACCAACCA 656  
QY 251 CAGATCAAGACACAAACCCCA-ACATACCTCACTCAGGATCTCTCAGCTTGAATCAG 309  
DB 655 CAACACACCCCAACAAACACACACACCAACCAACCAACCAACCAACCAACCAAC 596  
QY 310 CTCCAATCTGTCTGAAATTTACATCAACAAAC-CACCACCATAGTCTTCAACCAAC 368  
DB 595 ACAAACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 536  
QY 369 GAGTCAAGTCAAACTGCAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCA 428  
DB 535 NAACAAACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 476  
QY 429 CACAACCCGCAAGCCCACTTACAAACCAACGCGCAAAACCAACCAACCAACCA 488  
DB 475 CAACAAACAAACACACACACCAACCAACCAACCAACCAACCAACCAACCAAC 416  
QY 489 ATGATTTTCACTTCTGAAAGTGTAACTTTGACCTGCGAGCATATGCGCAACCA 548  
DB 415 ACAACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 356  
QY 549 CCTGCTGGCTTATCTGCAAAAGAAATACCAACAAAAACCAAGGAAGAAACCAAC 608  
DB 355 ACACACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 296  
QY 609 AGCTTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCAACCA 668  
DB 295 CACCAAAACCAAAACCAAAACCAACCAACCAACCAACCAACCAACCAACCA 236  
QY 669 AACCAAGGAGTACCCACCAACCAACGCGCCACAGAGGCGCAACCATCAACCA 728  
DB 235 ACACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 176  
QY 729 CAACATCAACAACTCACTGCTGCTACCAACCAACCAACCAACCAACCAACCA 773  
DB 175 CACCAACAAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCA 131

RESULT 6  
US-10-312-841-2/c  
; Sequence 2, Application US/10312841  
; Publication No. US20030186277A1  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC  
; FILE REFERENCE: E01/1208/WO  
; CURRENT APPLICATION NUMBER: US/10/312,841  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 2  
; SEQ ID NO 2  
; LENGTH: 3673778  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (379615)  
US-10-312-841-2  
Query Match  
9.3%; Score 85.8; DB 12; Length 3673778;

Best Local Similarity 47.0%; Pred. No. 2.8e-10;  
Matches 334; Conservative 0; Mismatches 372; Indels 5; Gaps 2;  
QY 214 CACACTAACCACTGCAATCATACAGATGCAAGCCAGATCAAGCAACCAACCCCAAC 273  
Db 2435804 CTCAACAAACCTTACAAACACGAAACCAAAATCTCCACTAACCAATAAAACCAACCAACC 2435745  
QY 274 ATACCTCACTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAAATTAATC 333  
Db 2435744 ATCCCGACAAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2435685  
QY 334 ACAACACCAACCACTAGCTTCAACCAACACGAGGAGTCAAGTCAAACTGCAACCCAC 393  
Db 2435684 CCAACAAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2435625  
QY 394 AACAGTCAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 453  
Db 2435624 AACAAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2435565  
QY 454 ACAACGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 513  
Db 2435564 AAAC-CTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2435506  
QY 514 CTTTGTACCTGAGCATGAGCAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 573  
Db 2435505 CAACAAACCTTACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 2435446  
QY 574 ACCAAACAAA---AACAGGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCT 629  
Db 2435445 CCCAACAAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2435386  
QY 630 TCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 689  
Db 2435385 CAACAAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2435326  
QY 690 CCAAGCCCAAGAGCCCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCAACCTGC 749  
Db 2435325 CACAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2435266  
QY 750 TCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 809  
Db 2435265 TAACAAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2435206  
QY 810 CTTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCAAACCAACCTCGAGCACCACATC 869  
Db 2435205 CCTCAACAAACCTTACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 2435146  
QY 870 AACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 920  
Db 2435145 CATCCCAACCAACCAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2435095

## RESULT 7

US-10-017-161-1981  
; Sequence 1981, Application US/10017161  
; Publication No. US20030143668A1  
; GENERAL INFORMATION:  
; APPLICANT: SUWA, MAKIKO  
; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: AKIYAMA, YUTAKA  
; APPLICANT: ABURATANI, HIROYUKI  
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
; FILE REFERENCE: 084335/0152  
; CURRENT APPLICATION NUMBER: US/10/017,161  
; CURRENT FILING DATE: 2002-12-18  
; PRIOR APPLICATION NUMBER: JP 2001/246789  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 2430  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1981  
; LENGTH: 1972  
; TYPE: DNA  
; ORGANISM: Homo sapiens

FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(1972)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (201)..(1772)  
; FEATURE:  
; NAME/KEY: modified base  
; LOCATION: (975)..(394)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
; FEATURE:  
; NAME/KEY: modified base  
; LOCATION: (1319)..(1328)  
; OTHER INFORMATION: a, t, c, g, unknown or other  
US-10-017-161-1981

Query Match 9.2%; Score 84.2; DB 12; Length 1972;  
Best Local Similarity 45.4%; Pred. No. 1.6e-11;  
Matches 350; Conservative 0; Mismatches 418; Indels 3; Gaps 3;

QY 128 ATCACTATTCATTTCTGGCAATGATAATCTCAACTTCTACTTATATATACAGCCATCAT 187  
Db 594 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 653  
QY 188 TTCAATAGCTCGGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 247  
Db 654 ATCAACCGCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 713  
QY 248 AGCCAGATCAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 307  
Db 714 ATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 773  
QY 308 TTCTCCAATCTGTGAAATTAATCAACCAACCAACCAACCAACCAACCAACCAACCAAC 366  
Db 774 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 833  
QY 367 AGGAGTCAAGTCAAAACCTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 425  
Db 834 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 893  
QY 426 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 485  
Db 894 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 953  
QY 486 ATATATGATTTTCACTTCGAGAGTTTAACTTTGTAGCTTGTAGCTATATGAGCAACATC 545  
Db 954 ACCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1013  
QY 546 CAACCTGTGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGGAAAGAAACCAACCA 605  
Db 1014 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1073  
QY 606 CCAAGCTTACAAACCAACCAACCTTCAAGCAACCAACCAACCAACCAACCAACCTCAACCA 665  
Db 1074 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1133  
QY 666 CTAACCAAGGAAGTACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 725  
Db 1134 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1193  
QY 726 AAACCAACCAACCAACCAACCTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAAC 785  
Db 1194 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1253  
QY 786 GTCAATGGAAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCT 845  
Db 1254 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1312  
QY 846 CCACCAACCTCGAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 896  
Db 1313 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1363



```
RESULT 8
US-10-311-455-1056/c
; Sequence 1056, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1056
; LENGTH: 16167
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 4003, 5529, 5534, 5553, 5570, 5573, 5577..5578, 5592, 5732
; OTHER INFORMATION: n is a or g or c or t
; NAME/KEY: unsure
; LOCATION: 5837..5838
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1056

Query Match      8.8%; Score 81.4; DB 12; Length 16167;
Best Local Similarity 46.6%; Pred. No. 2.5e-10;
Matches 235; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

Qy 314 AATCTGTCTGAAATTACATCAAAACCAACCACTACTAGTCTTCAACCAACCAAGAGTC 373
Db 5866 ATTCTTACTTTTCTCACACCAAAACGNNNAACAAACGAAAAAAGAAAAAAGAAAAA 5807
Qy 374 AGTCAAACTGCAACCCCAACAGCTCAAGCTTAAACAAACACACACACACCAACCA 433
Db 5806 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5747
Qy 434 CCCAGCAAGCCCACTACAAACCAACGCAAAACAAACCAACCAACCAACCAATATGAT 493
Db 5746 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5687
Qy 494 TTTCACTTGTGAAGTGTAACTTTGTACCTGCGAGCATATGCGACAACTCAACCTGCG 553
Db 5686 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5627
Qy 554 TGGGCTATCTGCAAGAAATACCAACCAAAACCAAGAAAGAAAGAAAGAAAGAAAGAA 613
Db 5566 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5507
Qy 674 AAGGAAGTACCAACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 733
Db 5506 AACCTTTTACCAACCAACCAATTAATCATCATCACTTAACTTAACTTAACTTAACTT 5447
Qy 734 ATCACTACTACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 793
Db 5446 TCTCCCCCAATTCCTCTCAACCTTTCTTATCAACCAACCTTAACTTAACTTAACTTAA 5387
Qy 794 GAAACCTTCCACTCAACCTCTCTCC 817
```

```
Db 5386 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5363

RESULT 9
US-10-240-485-82/c
; Sequence 82, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240.485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 82
; LENGTH: 16167
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (4003, 5529, 5534, 5553, 5570, 5573, 5577..5578, 5592, 5732)
; NAME/KEY: unsure
; LOCATION: (5837..)
US-10-240-485-82

Query Match      8.8%; Score 81.4; DB 12; Length 16167;
Best Local Similarity 46.6%; Pred. No. 2.5e-10;
Matches 235; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

Qy 314 AATCTGTCTGAAATTACATCAAAACCAACCACTACTAGTCTTCAACCAACCAAGAGTC 373
Db 5866 ATTCTTACTTTTCTCACACCAAAACGNNNAACAAACGAAAAAAGAAAAAAGAAAAA 5807
Qy 374 AGTCAAACTGCAACCCCAACAGCTCAAGCTTAAACAAACACACACACCAACCA 433
Db 5806 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5747
Qy 434 CCCAGCAAGCCCACTACAAACCAACGCAAAACAAACCAACCAACCAACCAATATGAT 493
Db 5746 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5687
Qy 494 TTTCACTTGTGAAGTGTAACTTTGTACCTGCGAGCATATGCGACAACTCAACCTGCG 553
Db 5686 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5627
Qy 554 TGGGCTATCTGCAAGAAATACCAACCAAAACCAAGAAAGAAAGAAAGAAAGAAAGAA 613
Db 5626 CGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5567
Qy 614 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 673
Db 5566 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5507
Qy 674 AAGGAAGTACCAACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 733
Db 5506 AACCTTTTACCAACCAACCAATTAATCATCATCACTTAACTTAACTTAACTTAACTT 5447
Qy 734 ATCACTACTACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 793
Db 5446 TCTCCCCCAATTCCTCTCAACCTTTCTTATCAACCAACCTTAACTTAACTTAACTTAA 5387
Qy 794 GAAACCTTCCACTCAACCTCTCTCC 817
```

QY 734 ATCACTACTACTCTCTCAACCAACACCAAGGAAATCTCAAACTCACAAGTCAATG 793  
 Db 5446 TCTCCCCCAATTCCTCAAACTCTCTATCAACCCCTAATAAATAAAAAAAAAAAAA 5387  
 QY 794 GAAACCTTCCACTCAACCTCTCTCC 817  
 Db 5386 AAAAAAAAAAAAAATAAATCTCTCTCC 5363

```

RESULT 10
US-10-311-455-378/c
; Sequence 378, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation of a Specific Gene
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 378
; LENGTH: 6270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-378

```

|    | Query Match           | 8.8%;   | Score 81;        | DB 12;    | Length 6270; |
|----|-----------------------|---|------------------|-----------|--------------|
|    | Best Local Similarity | 45.1%;  | pred. No. 2e-10; |           |              |
|    | Matches 345;          | Conservative 0;   | Mismatches 415;  | Indels 5; | Gaps 1;      |
| QY | 38                    | AAGACACTAGAAAGACCTGGGACACTCTCAATCATTTATTATTCATATCATCTCGGGCTTA | 97               |           |              |
| Db | 2056                  | AACATATCTTAAACACACGCGAACTCGGGTATAAAAACTTATTCCTTAAACCTTATCACT  | 1997             |           |              |
| QY | 98                    | TATAAGTTAAATCTTAAATCTGTAGCACAATCAATTCATTCCTGGCAATCATATTC      | 157              |           |              |
| Db | 1996                  | ACTCCCTTAAACCCCTCAAAAAACAAACCCCTAACCTACACTCAAAAAATCTTAAAAAAT  | 1937             |           |              |
| QY | 158                   | TCAACTTCATTATATATAGCCATCATTTCTATAGCCTGGGCAACCAACCAAGTCACA     | 217              |           |              |
| Db | 1936                  | AAAATCATAAAAAATACTTTCTCTTAAAAATAAATAAAACCAACCAAAAAAACCACCA    | 1877             |           |              |

|    |      |  |      |
|----|------|--|------|
| QY | 218  | CTAACCACTGCAATCATACAAGATGCAACAAGCCAGATCAAGACACAAACCCCAACATAC     | 277  |
| Db | 1876 | AAAAAAAAAGCTATCAAAAAAAAAAAAAATAAAACCCCACTAAAAATAAACACTAAAAAATAAC | 1817 |
| QY | 278  | CTCACTCAGGATCCTCAGCTTTGGAATCAGCTTTCTCC-----AATCTGTCTGAAATTACAT   | 332  |
| Db | 1816 | TTTCTCTCTCATCTCTTAACCTTAAACCTCAATCTTACATAAACCTTTTCAAAATCGAAAA    | 1757 |
| QY | 333  | CACAAACCAACCACTACTAGTTCTCAACACACGAGGAGTCAAGTCAAACTGCAACCCCA      | 392  |
| Db | 1756 | AAAAAAATTACTCAAAATAAAATATTACCTCTACGAAATTTTAAATCCCTTAACCTCC       | 1697 |
| QY | 393  | CACAGTCAAGCTTAAAAACAAACAAACCCAAACAAACCAACCAAGCCCACTACAA          | 452  |
| Db | 1696 | CACAAAAATAAAAAAACAATCCCAAAAAATCAACTTCGAAACATCTCCAAACCCCTAA       | 1637 |
| QY | 453  | AACAAGCCGAAAAACAACCAACCAATAATGTTTTCTATTCTCGAAGTGTTTA             | 512  |
| Db | 1636 | CTTAATCCCTAACCAAAAAATATCCCAACTACTTAATAACCAAAATCGCCTTTCTCT        | 1577 |

|    |      |   |      |
|----|------|---|------|
| QY | 513  | ACTTTGTAACCTGCAGCATATGCGAGCAACAAATCAACCTGTGGGTATCTGCAAAAGAA     | 572  |
| Db | 1576 | AAAACTTACATCTCTCCCTCTAAATAATATCTCTACCCCTCCCTAAATCACCCTCAAAAAA   | 1517 |
| QY | 573  | TACCAACCAAAAAACGAGGAAGAAAACCCACCAACGCTTACAAATAAAACCAACCTTCA     | 632  |
| Db | 1516 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAATCTACAAATAAAATAAAATAAAACA | 1457 |
| QY | 633  | AGACACCAAAAAAGATCTCAAACTCAAACTCAAAACCAAGGAAGTACCACCAACCA        | 692  |
| Db | 1456 | AAACAAAAATAAAAAAAAAAAAAACCAACAAAAATAATTAACAAAAAACTACAAAAACAA    | 1397 |
| QY | 693  | AGCCACAGAGAGCCAAACCATCAACACCCCAACCAACCAACATCACACTACACTGTCTCA    | 752  |
| Db | 1396 | ATCTTAAACAAAAAATCGAAAAAATAACAAACCCCAATCAAAAAATAAAAAATAAAACA     | 1337 |
| QY | 753  | CCAAACACACCAAGGAAATCCAAAACTCAAGTCAAATGGAAA                      | 797  |
| Db | 1336 | ACCGAAACACCCCAAAAAATACCTCAAAAAACAATAAAAAACGAAA                  | 1292 |

```

RESULT 11
US-10-240-453-20/c
; Sequence 20, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPNEROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation St
; TITLE OF INVENTION: With DNA Transcription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240.453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 20
; LENGTH: 6270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapi
US-10-240-453-20

```

|    | Query Match           | 8.8%   | Score 81;        | DB 12;          | Length 6270;      |
|----|-----------------------|--|------------------|-----------------|-------------------|
|    | Best Local Similarity | 45.1%  | Pred. No. 2e-10; |                 |                   |
|    | Matches 345;          | Conservative   | 0;               | Mismatches 415; | Indels 5; Gaps 1; |
| QY | 38                    | AAGACACTAGAAAAGACCTGGGACACTCTCAATCATTATTATTCATATCATTCGGGCTTA     | 97               |                 |                   |
| DB | 2056                  | AACATACTCTTAAACACACGCGAACTTCGGGTATAAACTTATTCCTTAAACACTATCACT     |                  |                 |                   |
| QY | 98                    | TATAAGTTTAAATCTTAAATCTGTAGCACAAATCACATTTATCCATTCCTGGCAATGATTAATC | 157              |                 |                   |
| DB | 1996                  | ACTCCCTTAAACCCCTCAAAAAACAAACCCCTAAACCTACACTCAAAAAATACTCTTAAAAAAT | 1937             |                 |                   |
| QY | 158                   | TCAACTTTCATTATAATTACAGCCATCATATTCATAGCCTTCGCGAACCCCAAGTCACA      | 217              |                 |                   |
| DB | 1936                  | AAARCTTACAAAAAATACTTCTCTTAAAAATAATAATAAAAACCCCAACCAAAAAAACCA     | 1877             |                 |                   |
| QY | 218                   | CTAACACTGCATCATACAAGATGCAACAAGCCAGATCAAGAACACCAACCCCAACAATAC     | 277              |                 |                   |

Db 1876 AAAAAAAGTATCAAAAAAATTAACCCGAACTAAATTAACACTAAAAAAGAC 1817  
QY 278 CTCACCTCAGAGTCTCAGCTTGGATGACCTTCTCC-----AATCTGTCTGAAATTTACAT 332  
Db 1816 TTTCTCTCTCACTCTAACTTAAACCTCAATCTACATATAAACTTTTTCATAAATCGAAAA 1757  
QY 333 CACAAACACACACCATCTAGCTTCAACAAACACGAGGATCAAGTCAAACTCGCAACCA 392  
Db 1756 AAAAAAATTTACTCAAAATTAATTTTACCTCTACGAAAAATTTAAATCCCTAACCTCC 1697  
QY 393 CAACAGTCAAGTCAAAAAACACAAACCAACCAACCAACCAACCAACCAACCAACCA 452  
Db 1696 CAACAAAAATTAATAAAACATTTCCCAAAACATCACTTCGAACTCTCCAAAACCTTAA 1637  
QY 453 AACAAGCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 512  
Db 1636 CTTAATCCCTAACCAAAAAATATCCCACTACTTAATAAACCAAAAAATCGCTTCTCT 1577  
QY 513 ACTTTGTACCTCGCAGCATATGAGCAACAATCCACCTGCTGGGCTATCTGCAAAAGAA 572  
Db 1576 AAAAACTTACTCTCTCCCTTAAATATATCTACCCCTCCCTAAATCAACCAAAAAA 1517  
QY 573 TACCAACAAAAACCAAGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCTTCA 632  
Db 1516 AAAAAAATTTACTCAAAATTAATTTTACCTCTAGCAAAATTTAAATCCCTTAACCTCC 1457  
QY 633 AGACAACCAAAAAAGATCTCAAACTTCAACCACTTAACCAACCAAGGAGTACCCACCA 692  
Db 1456 AAAAAAATTAATAAAACCAACCAAAAAATTAATAACCAAAAAAATCTACAAAAACAA 1397  
QY 693 AGCCACAGAGGCAACCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 752  
Db 1396 ATCTAAACAAAAATCGAAAAATCGAAAAATCGAAAAATCGAAAAATCGAAAAAC 1337  
QY 753 CCAACAAACACCAAGAAATCGAAAAATCGAAAAATCGAAAAATCGAAAAATCGAAAA 797  
Db 1336 ACGAAACACCCCAAAAAATCTCTCAAAACCAATTAATAAACGAAA 1292

## RESULT 12

US-10-239-676-12/c  
; Sequence 12, Application US/10239676  
; Publication No. US2003008260A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
; FILE REFERENCE: 5013.1003  
; CURRENT APPLICATION NUMBER: US/10/239,676  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: PCT/EP01/03968  
; DE 10019058.8  
; DE 10019173.8  
; DE 10032529.7  
; DE 10043826.1  
; PRIOR FILING DATE: 2001-04-06  
; 2000-04-06  
; 2000-04-07  
; 2000-06-30  
; 2000-09-01  
; NUMBER OF SEQ ID NOS: 228  
; SEQ ID NO 12  
; LENGTH: 6270  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-239-676-12

Query Match 8.8%; Score 81; DB 14; Length 6270;  
Best Local Similarity 45.1%; Pred. No. 2e-10;

Matches 345; Conservative 0; Mismatches 415; Indels 5; Gaps 1;  
QY 38 AAGACACTAGAAAAAGACCTGGGACACTCTCAATCAATTTATTATTATTCATATATCGGGCTTAA 97  
Db 2056 AACATACCTTAAACACCAACGCAACTCGCGTATATAAACTTTATTCCTAAACCTTATCACT 1997  
QY 98 TATAAGTTTAAATCTTAAATCTGTAGCACAATATCAATTTATCGATTTCTGGCAATGATATC 157  
Db 1996 ACTCCCTAAACCCCTCAAAAAACCAACCCCTAACCTTACACTCANAATAATCTCTAAAAAT 1937  
QY 158 TCAACTTCACTTATATAATTACAGCCATCATATTTCTAGCTCGGCAACCAACCAAGTCA 217  
Db 1936 AAAAACTCAAAAAAATTTCTTTTAAATAATTAATAAACCCCAACCAAAAAAACA 1877  
QY 218 CTAACAATCGAATCATACAAGATGCAACAGCCAGATCAAGAAACCAACCCCAACATAC 277  
Db 1876 AAAAAAAGCTATCAAAAAAATAAAACCACTAAATAAACAATAAACAATAAACAATAAACA 1817  
QY 278 CTCACCTCAGGATCCTCAGCTTGGATCAGCTTCTCC-----AATCTGTCTGAAATTTACAT 332  
Db 1816 TTTCTCTCTTCACTCTAACTTAAACCCCTCATTTCTACATAAAACCTTTTTCAAAAATCGAAA 1757  
QY 333 CACAAACCAACCAACCATCTAGCTTCAACCAACCAACGAGGATCAAGTCAAAACCTGCAACCCA 392  
Db 1756 AAAAAAATTTACTCAAAATTAATTTTACCTCTAGCAAAATTTAAATCCCTTAACCTCC 1697  
QY 393 CAACAGTCAAGACTAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 452  
Db 1696 CAACAAAAATTAATAAAACATTTCCCAAAACATCAACTTCGAAACATCTCCAAAACCTTAA 1637  
QY 453 AACACGCGCAAAACCAACCAACCAACCAACCAACCAATTAATGATTTTTCATCTCGAAGTGTTA 512  
Db 1636 CTTAATCCCTAACCAAAAAATATCCCACTACTTAATAAACAACCAAAAAATCGCTTCTCT 1577  
QY 513 ACTTTGTACCTCGCAGCATATGAGCAACAATCGAACTCTGCTGGGCTATCTGCAAAAGAA 572  
Db 1576 AAAAACTTACTCTCTCCCTTAAATATATCTACCCCTCCCTAAATCAACCCCAAAAAA 1517  
QY 573 TACCAACAAAAACCAAGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCTTCA 632  
Db 1516 AAAAAAATTAATAAAACCAAAAAAATTAATAACCAAAAAAATTAATAAACA 1457  
QY 633 AGACAACCAAAAAAGATCTCAAACTTCAACCACTTAACCAACCAAGGAGTACCCACCA 692  
Db 1456 AAAAAAATTAATAAAACCAACCAAAAAATTAATAACCAAAAAAATCTACAAAAACAA 1397  
QY 693 AGCCACAGAGGCAACCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 752  
Db 1396 ATCTAAACAAAAATCGAAAAATCGAAAAATCGAAAAATCGAAAAATCGAAAAAC 1337  
QY 753 CCAACAAACACCAAGAAATCGAAAAATCGAAAAATCGAAAAATCGAAAAATCGAAAA 797  
Db 1336 ACCGAAACACCCCAAAAAATCTCTCAAAACCAATTAATAAACGAAA 1292

## RESULT 13

US-10-311-455-1670/c  
; Sequence 1670, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1

```
/ PRIOR FILING DATE: 2000-09-01
/ NUMBER OF SEQ ID NOS: 2424
/ SEQ ID NO 1670
/ LENGTH: 6668
/ TYPE: DNA
/ FEATURE:
/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 1936
/ OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1670

Query Match      8.7%; Score 80; DB 12; Length 6668;
Best Local Similarity 48.9%; Pred. No. 3.7e-10;
Matches 215; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 358 AACACACACAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGCTATAAAACACACAC 417
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 418 AACACCCAAACACACCCAGCAGCCCACTACAAAACAAACGCCCCAAACAAACCCACAAA 477
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 CAACCCCAATGATTTTCACTTCGAGTGTGTTTAACTTTGTACCTGCGCATATGCGAG 537
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 CAACCAATCCACTGCTGGCTATCTGCAAAAGAGATACCAACCAAAACCCAGGAAAGAA 597
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 598 AACCAACCAAGCCTACAAAAACCAACCTTCAAGCAACCAACCAAAAGATCTCAAAACC 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 658 TCAACCACTAACCAAGGAGTACCCACCAACCCAGCCACAGAGCCCAACCATCAA 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 718 CACCAACCAAAACAAACATCACTACACTGCTCACCACCAACCAACCAAGGAATCCAAA 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 778 ACTCAAGTCAATGGAAA 797
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2931 AAAACCAAAAAAAACAAA 2912
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-864-761-20241
/ Sequence 20241, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aecmics-x-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
```

```
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 20241
/ LENGTH: 1635
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC006547.9
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864-761-20241

Query Match      8.5%; Score 78.2; DB 9; Length 1635;
Best Local Similarity 46.6%; Pred. No. 5.3e-10;
Matches 354; Conservative 0; Mismatches 398; Indels 7; Gaps 3;

QY 128 ATCACTTATTCATTCTGGCAATGATTAATCTCAACTTTCACCTTAATTTACAGCCATCATA 187
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 377 ACCACCAATACCATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATCA 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 TTCTAGCTCTGGCAAAACCCAAAGTCACTAACTCACTCACTCACTCACTCACTCACTCACT 247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 437 ACCATCACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 496
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 248 AGCCAGATCAAGACACACACCCCAACATCACTCACTCACTCACTCACTCACTCACTCACT 307
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 497 ATCACTCACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 308 TTCTCCAATCTGTCTGAAATTTACATCAAAACCAACCAACCACTAGCTTCAACCAACCA 367
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 557 ATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 368 GGAGTCAAGTCAAAACCTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 617 ACCACCAACCATCACTGCGCACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 676
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 428 ACACACCCAGCAGCCCACTACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 487
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 677 ACTGCCACACACACACACACACACAT-CATTCACACACACACACACACACACAC 735  
Qy 488 AATGATTTTCTCTTGAAGTGTTTAACTTTGTACCTGTCAGCATATGACGACACATCCA 547  
Db 736 CATCACTACT 790  
Qy 548 ACTGCTGGGTCTCTGAAAGATATACAAACAAACACAGGAAAGAAACACACACAC 607  
Db 791 ACCAC 850  
Qy 608 AAGCTTACAAACACACACCTTCAAGACACACACACACACACACACACACACAC 667  
Db 851 ACTAC 910  
Qy 668 AAACCAAGGAAGTAC 727  
Db 911 ACCAC 970  
Qy 728 ACAACATCACAAC 787  
Db 971 ACCACATCACAAC 1030  
Qy 788 CAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGC-CCTTCTCAAGTCTC 846  
Db 1031 ACCATCATCACAAC 1090  
Qy 847 CACAACATCGAGCACCACCATCACAACCTCTCTCCACC 885  
Db 1091 ACTAC 1129

## RESULT 15

US-09-864-761-3471  
; Sequence 3471, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemica-X-1  
; CURRENT FILING DATE: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 3471  
; LENGTH: 1973  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURES:  
; OTHER INFORMATION: MAP TO AC006547.9  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12  
; US-09-864-761-3471

Query Match 8.5%; Score 78.2; DB 9; Length 1973;

Best Local Similarity 45.6%; Pred. No. 5.9e-10;

Matches 354; Conservative 0; Mismatches 398; Indels 7; Gaps 3;

Qy 128 ATCACTTATCCATTCTGGCAATGATTAATCTCAACTTCACTTAAATTACAGCCATCATA 187  
Db 531 ACCACCATACCATCACTTACCACCATCAACACACACACCATCACTACCATCACCACCATC 590  
Qy 188 TTGATAGCTGGCAAAACCAAGTCACTAACAATGCACTGCAATGCAATGCAATGCAATG 247  
Db 591 ACCATCACTATACACACACACACACACACCATCACTACCATCACTACCATCACTACCATC 650  
Qy 248 AGCCAGATCAAGACACACACACACACACATCACTCACTCAGGATCCTCAGTTGGATCAGC 307  
Db 651 ATCACTATCACTACACACACACACACACCATCACTACCATCACTACCATCACTACCATC 710  
Qy 308 TTCTCCAACTGTCTGAAATTTACATCAAAACACACACACATCACTAGTTTCAACACACCA 367  
Db 711 ATCAACACACACACACACACACACACATCACTACCATCACTACCATCACTACCATC 770  
Qy 368 GGAGTCAAGTCAAACTTGCAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 427  
Db 771 ACCACACCATCACTGCGCACACACACACACACCATCACTACCATCACTACCATCACTACCATC 830  
Qy 428 AC 487  
Db 831 ACTGCAAC 889  
Qy 488 AATGATTTTCACTTCTGAAAGTGTTTAACTTTGTACCTGTCAGCATATGACGACACATCCA 547  
Db 890 CATCACTAC 944  
Qy 548 ACTGCTGGGTATCTGCAAAAGATATACCAAAACCAAAACACAGGAAAGAAACACACACAC 607  
Db 945 ACCACACACACATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATC 1004  
Qy 608 AAGCTTACAAACACACACCTTCAAGACACACACACACACACACACACACACACACAC 667  
Db 1005 ACTTACACACACACACACATCACTACCATCACTACCATCACTACCATCACTACCATC 1064  
Qy 668 AAACCAAGGAAGTATCCACACACACACACACACACACACACACACACACACACAC 727  
Db 1065 ACCACACCATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATC 1124

[illegible]

Search completed: October 30, 2003, 03:23:48  
Job time : 319.043 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 21:37:29 ; Search time 71.4618 Seconds  
(without alignments)  
5682.373 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920  
Sequence: 1 tgcacacatgtccaaaacaa.....gtagtattataaaaaaaaaa 920

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A-COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B-COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A-COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B-COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS-COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 920   | 100.0       | 920    | 2  | US-08-467-963C-7  |
| 2          | 920   | 100.0       | 920    | 2  | US-08-838-189D-7  |
| 3          | 920   | 100.0       | 920    | 3  | US-08-852-344D-7  |
| 4          | 920   | 100.0       | 920    | 3  | US-08-344-639E-7  |
| 5          | 920   | 100.0       | 920    | 3  | US-08-467-969A-7  |
| 6          | 920   | 100.0       | 920    | 3  | US-08-467-961A-7  |
| 7          | 920   | 100.0       | 920    | 3  | US-08-001-554A-7  |
| 8          | 894   | 97.2        | 894    | 2  | US-08-467-963C-28 |
| 9          | 894   | 97.2        | 894    | 2  | US-08-838-189D-28 |
| 10         | 894   | 97.2        | 894    | 3  | US-08-852-344D-28 |
| 11         | 894   | 97.2        | 894    | 3  | US-08-344-639E-28 |
| 12         | 853.4 | 92.8        | 15222  | 2  | US-08-801-898A-23 |
| 13         | 853.4 | 92.8        | 15222  | 2  | US-08-962-690-12  |
| 14         | 853.4 | 92.8        | 15223  | 2  | US-08-892-403A-1  |
| 15         | 853.4 | 92.8        | 15223  | 2  | US-08-720-132-1   |
| 16         | 441.4 | 48.0        | 15225  | 2  | US-08-892-403A-2  |
| 17         | 212.8 | 23.1        | 1050   | 4  | US-08-836-501-76  |
| 18         | 212.8 | 23.1        | 1050   | 4  | US-09-626-830-76  |
| 19         | 210.8 | 22.9        | 303    | 3  | US-08-721-979A-1  |
| 20         | 210.8 | 22.9        | 303    | 3  | US-08-836-504A-1  |
| 21         | 210.8 | 22.9        | 303    | 3  | US-08-836-501-1   |
| 22         | 210.8 | 22.9        | 303    | 4  | US-09-654-289-1   |
| 23         | 210.8 | 22.9        | 303    | 4  | US-09-582-876-1   |
| 24         | 210.8 | 22.9        | 303    | 4  | US-09-626-830-1   |
| 25         | 209.6 | 22.8        | 1071   | 3  | US-08-836-501-77  |
| 26         | 209.6 | 22.8        | 1071   | 4  | US-09-626-830-77  |
| 27         | 207.6 | 22.6        | 303    | 3  | US-08-721-979A-3  |

28 207.6 22.6 303 3 US-08-836-504A-3  
29 207.6 22.6 303 3 US-08-836-501-3  
30 207.6 22.6 303 3 US-09-654-289-3  
31 207.6 22.6 303 4 US-09-582-876-3  
32 207.6 22.6 303 4 US-09-582-876-3  
33 201.2 21.9 303 3 US-08-721-979A-14  
34 201.2 21.9 303 3 US-08-836-501-14  
35 201.2 21.9 303 4 US-09-654-289-14  
36 201.2 21.9 303 4 US-09-582-876-14  
37 201.2 21.9 303 4 US-09-626-830-14  
38 137 14.9 840 5 PCT-US91-08177-12  
39 137 14.9 7323 5 PCT-US91-08177-1  
40 128.6 14.0 183 3 US-08-721-979A-23  
41 128.6 14.0 183 3 US-08-836-501-23  
42 128.6 14.0 183 4 US-09-654-289-23  
43 128.6 14.0 183 4 US-09-582-876-23  
44 128.6 14.0 183 4 US-09-626-830-23  
45 126 13.7 303 3 US-08-721-979A-2

## ALIGNMENTS

### RESULT 1

US-08-467-963C-7

; Sequence 7, Application US/08467963C

; Patent No. 5968776

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michel H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHYN, Mary E

; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST

; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6th Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1K7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/467,963C

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/838,189

; FILING DATE: 16-APR-1997

; APPLICATION NUMBER: US 08/001,554

; FILING DATE: 06-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9200117.1

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-474 MIs:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 920 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-467-963C-7









|  |     |  |     |
|--|-----|--|-----|
| Db   | 781 | CACAAGTCAAAATGGAAACCTTCCACTCAACCTCTCCGAGGCAATCTTAAGCCCTTCTCA     | 840 |
| Qy   | 841 | AGTCTCCACAACATCCGAGCAGCCCATCAACAACCTCATCTCCACCCCAACAACACGCGCA    | 900 |
| Db   | 841 | AGTCTCCACAACATCCGAGCAGCCCATCAACAACCTCATCTCCACCCCAACAACACGCGCA    | 900 |
| Qy   | 901 | GTAGTTATTAAAAA920  |     |
| Db   | 901 | GTAGTTATTAAAAA920  |     |
| RESULT 7   |     |  |     |
| US-08-001-554A-7   |     |  |     |
| ; Sequence 7, Application US/08001554A                       |     |  |     |
| ; Patent No. 6225091   |     |  |     |
| ; GENERAL INFORMATION:                                       |     |  |     |
| ; APPLICANT: Klein, Michel H                                 |     |  |     |
| ; APPLICANT: Du, Run-Pan                                     |     |  |     |
| ; APPLICANT: Ewasysdyn, Mary E                               |     |  |     |
| ; TITLE OF INVENTION: Chimeric Immunogens                    |     |  |     |
| ; NUMBER OF SEQUENCES: 21                                    |     |  |     |
| ; CORRESPONDENCE ADDRESS:                                    |     |  |     |
| ; ADDRESSEE: Sim & McBurney                                  |     |  |     |
| ; STREET: 330 University Avenue, 6th Floor                   |     |  |     |
| ; CITY: Toronto  |     |  |     |
| ; STATE: Ontario   |     |  |     |
| ; COUNTRY: Canada  |     |  |     |
| ; ZIP: M5G 1R7   |     |  |     |
| ; COMPUTER READABLE FORM:                                    |     |  |     |
| ; MEDIUM TYPE: Floppy disk                                   |     |  |     |
| ; COMPUTER: IBM PC compatible                                |     |  |     |
| ; OPERATING SYSTEM: PC-DOS/MS-DOS                            |     |  |     |
| ; SOFTWARE: PatentIn Release #1.0, Version #1.25             |     |  |     |
| ; CURRENT APPLICATION DATA:                                  |     |  |     |
| ; APPLICATION NUMBER: US/08/001,554A                         |     |  |     |
| ; FILING DATE: 06-JAN-1993                                   |     |  |     |
| ; CLASSIFICATION: 435  |     |  |     |
| ; ATTORNEY/AGENT INFORMATION:                                |     |  |     |
| ; NAME: Stewart, Michael I                                   |     |  |     |
| ; REGISTRATION NUMBER: 24,973                                |     |  |     |
| ; REFERENCE/DOCKET NUMBER: 1038-286                          |     |  |     |
| ; TELECOMMUNICATION INFORMATION:                             |     |  |     |
| ; TELEPHONE: (416) 595-1155                                  |     |  |     |
| ; TELEFAX: (416) 595-1163                                    |     |  |     |
| ; TELEX: 065-24567 SIMBAS                                    |     |  |     |
| ; INFORMATION FOR SEQ ID NO: 7:                              |     |  |     |
| ; SEQUENCE CHARACTERISTICS:                                  |     |  |     |
| ; LENGTH: 920 base pairs                                     |     |  |     |
| ; TYPE: nucleic acid   |     |  |     |
| ; STRANDEDNESS: single                                       |     |  |     |
| ; TOPOLOGY: linear   |     |  |     |
| ; MOLECULE TYPE: DNA (genomic)                               |     |  |     |
| US-08-001-554A-7   |     |  |     |
| Query Match 100.0%; Score 920; DB 3; Length 920;             |     |  |     |
| Best Local Similarity 100.0%; Pred. No. 8.8e-243;            |     |  |     |
| Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |     |  |     |
| Qy   | 1   | TGCAAAACATCTCCAAAACAGAGCCACCAACGCGCTTAGACACTAGAAAGACCTGGGA       | 60  |
| Db   | 1   | TGCAAAACATCTCCAAAACAGAGCCACCAACGCGCTTAGACACTAGAAAGACCTGGGA       | 60  |
| Qy   | 61  | CACCTCTCAATTCATTATTATTCATATCATCTCGGCGCTTATATAAGTTAAATCTTAAATCTGT | 120 |
| Db   | 61  | CACCTCTCAATTCATTATTATTCATATCATCTCGGCGCTTATATAAGTTAAATCTTAAATCTGT | 120 |
| Qy   | 121 | AGCACAACATCACATTAATCCATTCTGGCAATGTAATCTCAACTTCACTTATAATTACAGC    | 180 |
| Db   | 121 | AGCACAACATCACATTAATCCATTCTGGCAATGTAATCTCAACTTCACTTATAATTACAGC    | 180 |
| Qy   | 181 | CATCATATTATATAGCTTCGGCAACCAACCAAGTCACACTAACTGCATATCATACAAGA      | 240 |
| Db   | 181 | CATCATATTATATAGCTTCGGCAACCAACCAAGTCACACTAACTGCATATCATACAAGA      | 240 |





[illegible]

```

RESULT 9
US-08-838-189D-28
; Sequence 28, Application US/08838189D
; Patent No. 5993169
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EMASYSHN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,189D  
FILING DATE: 16-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-838-189D-28

Query Match 97.2%; Score 894; DB 2; Length 894;  
Best Local Similarity 100.0%; Prod. No. 1.2e-235;  
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| QY | 8   | ATGTCCAAAAACAAGGACCAACGACCGCTAAGACACTAGAAAAAGACCTGGGACACTCTC    | 67  |
| Db | 1   | ATGTCCAAAAACAAGGACCAACGACCGCTAAGACACTAGAAAAAGACCTGGGACACTCTC    | 60  |
| QY | 68  | AATCATTTATTTATTCATATCATCTGGGCTTATATAGTTTAAATCTTAAATCTGTAGCAAA   | 127 |
| Db | 61  | AATCATTTATTTATTCATATCATCTGGGCTTATATAGTTTAAATCTTAAATCTGTAGCAAA   | 120 |
| QY | 128 | ATCACATTATTCCTTCTGGCAATGATTAATCTCAACTTTCACATTATAATTACAGCCATCAT  | 187 |
| Db | 121 | ATCACATTATTCCTTCTGGCAATGATTAATCTCAACTTTCACATTATAATTACAGCCATCAT  | 180 |
| QY | 188 | TTCATAGCTCGGCAAAACCAAAAGTCACACTAACTGCAATCATATAAGATGCAACA        | 247 |
| Db | 181 | TTCATAGCTCGGCAAAACCAAAAGTCACACTAACTGCAATCATATAAGATGCAACA        | 240 |
| QY | 248 | AGCCAGATCAAGAACACAAACCCCAACATACCTCACTCAGGATCCTCAGCTTTGGAAATCAGC | 307 |
| Db | 241 | AGCCAGATCAAGAACACAAACCCCAACATACCTCACTCAGGATCCTCAGCTTTGGAAATCAGC | 300 |
| QY | 308 | TTCTCCAAATCTGTCTGAAATTTACATCAACAACCAACCACTACTACTAGCTTTCACCAACCA | 367 |
| Db | 301 | TTCTCCAAATCTGTCTGAAATTTACATCAACAACCAACCACTACTACTAGCTTTCACCAACCA | 360 |
| QY | 368 | GGAGTCAAGTCAAAACCTTGCAACCCCAACAGTCAAGACTTAAAAACAAACAACACCCAA    | 427 |
| Db | 361 | GGAGTCAAGTCAAAACCTTGCAACCCCAACAGTCAAGACTTAAAAACAAACAACACCCAA    | 420 |
| QY | 428 | ACACACCCGAGGAGCCCACTACAAAACAAGCCGCAAAACAACCAACGAAACCAACCCCAAT   | 487 |
| Db | 421 | ACACACCCGAGGAGCCCACTACAAAACAAGCCGCAAAACAACCAACGAAACCAACCCCAAT   | 480 |
| QY | 488 | AATGATTTTTCACCTTCSAAGTGTTTAACTTTGTACCCCTGCGCATATGCAAGCAAACTCCA  | 547 |
| Db | 481 | AATGATTTTTCACCTTCSAAGTGTTTAACTTTGTACCCCTGCGCATATGCAAGCAAACTCCA  | 540 |
| QY | 548 | ACCTGCTGGGCTATCTGCAAAAGAATACCAAAACAAAAACCCAGGAAAGAAAAACCAACC    | 607 |
| Db | 541 | ACCTGCTGGGCTATCTGCAAAAGAATACCAAAACAAAAACCCAGGAAAGAAAAACCAACC    | 600 |
| QY | 608 | AAGCCTACAAAAAAACCAACCTTCAAGACAAACCAAAAAAGATCTCAAACTCAAAACCACT   | 667 |
| Db | 601 | AAGCCTACAAAAAAACCAACCTTCAAGACAAACCAAAAAAGATCTCAAACTCAAAACCACT   | 660 |
| QY | 668 | AAACCAAGGAAGTACCACCAACCAAGCCCAAGAGGCAACCAACCTCAACCAACCAACCAAA   | 727 |
| Db | 661 | AAACCAAGGAAGTACCACCAACCAAGCCCAAGAGGCAACCAACCTCAACCAACCAACCAAA   | 720 |
| QY | 728 | ACAAACATCAACTACACTGCTCACCAACCAACCAAGGAAATCCAAAACTCACAAGT        | 787 |
| Db | 721 | ACAAACATCAACTACACTGCTCACCAACCAACCAAGGAAATCCAAAACTCACAAGT        | 780 |
| QY | 788 | CAAAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC      | 847 |
| Db | 781 | CAAAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC      | 840 |
| QY | 848 | ACAACATCCGAGCAACCCATCACACCTCATCTCCACCAACCAACCAAGCCGAG           | 901 |
| Db | 841 | ACAACATCCGAGCAACCCATCACACCTCATCTCCACCAACCAACCAAGCCGAG           | 894 |

RESULT 10

US-08-852-344D-28

; Sequence 28, Application US/08852344D

; Patent No. 6017539

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michel H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHIN, Mary E

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION

; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY

; TITLE OF INVENTION: SYNCYTIAL VIRUS

; NUMBER OF SEQUENCES: 37

```

RESULT 10
US-08-852-344D-28
; Sequence 28, Application US/08852344D
; Patent No. 6017539
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EWASYSHYN, Mary E
; TITLE OF INVENTION: CHIMERIC PROTEIN
; TITLE OF INVENTION: AGAINST PARASITIC
; TITLE OF INVENTION: SYNCTYAL VIRUS
; NUMBER OF SEQUENCES: 37

```

[illegible]

; INFORMATION FOR SEQ ID NO: 28:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 894 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-344-639E-28

Query Match 97.2%; Score 894; DB 3; Length 894;

Best Local Similarity 100.0%; Pred. No. 1.2e-235;

Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      8 ATCTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGACATCTTC 67
DB      1 ATCTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGACATCTTC 60
QY      68 AATCATTTATTATTTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGTAGCACAA 127
DB      61 AATCATTTATTATTTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGTAGCACAA 120
QY      128 ATCAGATTATTCATTTCTGGCAATGATAATCTCACTTCACTTATTAATTACAGCCATCAT 187
DB      121 ATCAGATTATTCATTTCTGGCAATGATAATCTCACTTCACTTATTAATTACAGCCATCAT 180
QY      188 TTCTATAGCTCGGCAACCAACCAAGTCACTACTAACCTGCAATCTATACAAAGATGCAACA 247
DB      181 TTCTATAGCTCGGCAACCAACCAAGTCACTACTAACCTGCAATCTATACAAAGATGCAACA 240
QY      248 AGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCTCAGCTTGGAAATCAGC 307
DB      241 AGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCTCAGCTTGGAAATCAGC 300
QY      308 TTCTCCAAATCTGTGAAATTTAGATCAACCAACCAAGTCACTACTAGCTTCAACCAACCA 367
DB      301 TTCTCCAAATCTGTGAAATTTAGATCAACCAACCAAGTCACTACTAGCTTCAACCAACCA 360
QY      368 GGAGTCAAGTCAACCTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 427
DB      361 GGAGTCAAGTCAACCTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
QY      428 ACACAAACCCAGAGCCCACTACAAAACCAACGCAACCAACCAACCAACCAACCAACCAAC 487
DB      421 ACACAAACCCAGAGCCCACTACAAAACCAACGCAACCAACCAACCAACCAACCAACCAAC 480
QY      488 AATGATTTTCACTTGGAAAGTGTAACTTTGTACCTGCGAGCATATGCGAGCAACATCCA 547
DB      481 AATGATTTTCACTTGGAAAGTGTAACTTTGTACCTGCGAGCATATGCGAGCAACATCCA 540
QY      548 ACCTGCTGGGCTATCTGCAAGGAATACCAACCAACCAACCAACCAACCAACCAACCAAC 607
DB      541 ACCTGCTGGGCTATCTGCAAGGAATACCAACCAACCAACCAACCAACCAACCAACCAAC 600
QY      608 AAGCTTACAAAACCAACCACTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAAC 667
DB      601 AAGCTTACAAAACCAACCACTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAAC 660
QY      668 AAACCAAGGAAGTATCCCAACCAACCAAGGCAACCAACCAACCAACCAACCAACCAACCA 727
DB      661 AAACCAAGGAAGTATCCCAACCAACCAAGGCAACCAACCAACCAACCAACCAACCAACCA 720
QY      728 ACRAACATCAACTACACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 787
DB      721 ACRAACATCAACTACACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 780
QY      788 CAATGGAAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCC 847
DB      781 CAATGGAAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCC 840
QY      848 ACAACATCCGAGCAACCAACCTCAATCTCCACCAACCAACCAACCAACCAACCAACCAAC 901
DB      841 ACAACATCCGAGCAACCAACCTCAATCTCCACCAACCAACCAACCAACCAACCAACCAAC 894
```

RESULT 12

US-08-801-898A-23

; Sequence 23, Application US/08801898A

; Patent No. 5998602

; GENERAL INFORMATION:

; APPLICANT: Torrence, Paul F.

; APPLICANT: Silverman, Robert H.

; APPLICANT: Cirino, Nick M.

; APPLICANT: Li, Guiying

; APPLICANT: Xiao, Wei

; TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE

; TITLE OF INVENTION: OLIGONUCLEOTIDES EFFECTIVE TO TREAT RSV INFECTIONS

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/801,898A

; FILING DATE: 18-FEB-1997

; CLASSIFICATION: 514

; NAME: Poissant, Brian M.

; REGISTRATION NUMBER: 28,462

; REFERENCE/DOCKET NUMBER: 8656-009

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-9741

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15222 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; FEATURE:

; NAME/KEY: RSV-A2

; LOCATION: 1...15222

; OTHER INFORMATION:

US-08-801-898A-23

Query Match 92.8%; Score 853.4; DB 2; Length 15222;

Best Local Similarity 95.5%; Pred. No. 5.4e-224;

Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY      1 TGCACACATGTCACAAAACAAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60
DB      4681 TGCACACATGTCACAAAACAAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 4740
QY      61 CACTCTCAATCAATTTATTATTATCATATCATCGTCTTATATAAGTTAAATCTTAAATCTGT 120
DB      4741 CACTCTCAATCAATTTATTATTATCATATCATCGTCTTATATAAGTTAAATCTTAAATCTGT 4800
QY      121 AGCACAATACATTTATCCATTTCTGGCAATGATAATCTCAACTTCACTTATTAATTACAGC 180
DB      4801 AGCACAATACATTTATCCATTTCTGGCAATGATAATCTCAACTTCACTTATTAATTACAGC 4860
QY      181 CATCATATTATAGCTTCGGCAACCAACCAAGTCACTAACTGAACCTGCAATCATACAAG 240
DB      4861 CATCATATTATAGCTTCGGCAACCAACCAAGTCACTAACTGCAATCATACAAG 4920
QY      241 TGCACAAGCCAGATCAAGAACCAACCCCAACATACCTCACCAGAAATCTTCAGCTTGG 300
DB      4921 TGCACAAGCCAGATCAAGAACCAACCCCAACATACCTCACCAGAAATCTTCAGCTTGG 4980
```



|  |  |      |
|--|--|------|
| 4922   | TGCAACAGCCGATCAAGAACACACACCCCAATACCTCACCAGCAATCCTCAGCTGG       | 498  |
| 301  | AATCAGCTTCTCCAAATCTGCTGAAATTAATACACAAACCAACCAATCTAGCTTCAAC     | 360  |
| 4982   | AATCAGTCCCTCTAAATCCTGCTGAAATTAATACACAAATCACCACCATCTAGCTTCAAC   | 5041 |
| 361  | AACACAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTTAAACACACACAAAC      | 420  |
| 5042   | AACACAGGAGTCAAGTCAAACTGCAACCCCTGCAATCCAAACAGTCAAGCAGCAACACAAAC | 5101 |
| 421  | AACCCAAACACAAACCCAGCAAGCCCTACTCAAAACAAACGCGCAAAACCAACCAACAA    | 480  |
| 5102   | AATCTAAACACAAACCCAGCAAGCCCAACCAAAACAAACGCGCAAAACCAACCAACAA     | 5161 |
| 481  | ACCCAATAATGATTTTTCACCTTGAAGTGTGTTAACTTTGTATCCCTGTCAGCATATG     | 540  |
| 5162   | ACCCAATAATGATTTTTCACCTTGAAGTGTGTTAACTTTGTATCCCTGTCAGCATATG     | 5221 |
| 541  | CAATCCAACTGCTGGCTATCTGCAAAAGAAATACCAAAACAAAAACCCAGGAAAGAAAC    | 600  |
| 5222   | CAATCCAACTGCTGGCTATCTGCAAAAGAAATACCAAAACAAAAACCCAGGAAAGAAAC    | 5281 |
| 601  | CACCAACAGCCTTACAAAAAACAACCTTCAAGACAAACCAAAAGAGATCTCAAACTCTCA   | 660  |
| 5282   | CATACCAAGCCCAAAAAAACAACCTTCAAGACAAACCAAAAGAGATCTCAAACTCTCA     | 5341 |
| 661  | AACCACTAAACCAAGGAAGTACCCACCAACCAAGCCCAAGAGAGAGCAACCACTCAAC     | 720  |
| 5342   | AACCACTAAATCAAGGAAGTACCCACCAACCAAGCCCAAGAGAGAGCAACCACTCAAC     | 5401 |
| 721  | CACCAAAACAAACATCAACACTACACTGCTCAACCAACCAACCAACCAAGAAATC        | 780  |
| 5402   | CACCAAAACAAACATCAACACTACACTGCTCAACCAACCAACCAACCAAGAAATC        | 5461 |
| 781  | CAGAAGTCAATGGAACCTTCCACTCAACTCTCTCCGAGGCAATCTAAGCCCTTCTCA      | 840  |
| 5462   | CAGAAGTCAATGGAACCTTCCACTCAACTCTCTCCGAGGCAATCTAAGCCCTTCTCA      | 5521 |
| 841  | AGTCTCCACAAACATCCGAGCAACCATCACAAACCTCTCTCCACCCCAACCAACGCA      | 900  |
| 5522   | AGTCTCTACAACTCCGAGTACCATCACAACTTCTCTCCACCCCAACCAACGCA          | 5581 |
| 901  | GTAGTTATTAATAAAAAA 919   |      |
| 5582   | GTAGTTACTTAAAAACATA 5600                                       |      |
| RESULT 15  |  |      |
| US-08-720-132-1  |  |      |
| ; Sequence 1, Application US/08720132                                  |  |      |
| ; Patent No. 6264957   |  |      |
| ; GENERAL INFORMATION:   |  |      |
| ; APPLICANT: Collins, Peter L.   |  |      |
| ; TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY             |  |      |
| ; TITLE OF INVENTION: SYNCYTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES |  |      |
| ; NUMBER OF SEQUENCES: 1   |  |      |
| ; CORRESPONDENCE ADDRESS:  |  |      |
| ; ADDRESSEE: Townsend and Townsend and Crew LLP                        |  |      |
| ; STREET: Two Embarcadero Center, 8th Floor                            |  |      |
| ; CITY: San Francisco  |  |      |
| ; STATE: CA  |  |      |
| ; COUNTRY: U.S.A.  |  |      |
| ; ZIP: 94111-3834  |  |      |
| ; COMPUTER READABLE FORM:  |  |      |
| ; MEDIUM TYPE: Floppy disk   |  |      |
| ; COMPUTER: IBM PC compatible  |  |      |
| ; OPERATING SYSTEM: PC-DOS/MS-DOS                                      |  |      |
| ; SOFTWARE: PatentIn Release #1.0, Version #1.25                       |  |      |
| ; CURRENT APPLICATION DATA:  |  |      |
| ; APPLICATION NUMBER: US/08/720,132                                    |  |      |
| ; FILING DATE: 27-SEP-1996   |  |      |
| ; CLASSIFICATION: 424  |  |      |
| ; PRIOR APPLICATION DATA:  |  |      |

; APPLICATION NUMBER: US 60/007,083  
; FILING DATE: 27-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 15280-250-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-467-9600  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15223 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other  
US-08-720-132-1

Query Match 92.8%; Score 853.4; DB 3; Length 15223;

Best Local Similarity 95.5%; Pred. No. 5.4e-224;

Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

|    |      |   |      |
|----|------|---|------|
| Qy | 1    | TGCAAAACATGTCACAAACAGGACCAACGCGCTAAGACACTAGAAAAGACCTGGGA        | 60   |
| Db | 4682 | TGCAAAACATGTCACAAACAGGACCAACGCGCTAAGACACTAGAAAAGACCTGGGA        | 4741 |
| Qy | 61   | CACCTCTCAATCATTTATTATTCATATCATCGGGCTATATAAGTTAAATCTTAAATCTGT    | 120  |
| Db | 4742 | CACCTCTCAATCATTTATTATTCATATCATCGGGCTATATAAGTTAAATCTTAAATCTGT    | 4801 |
| Qy | 121  | AGCAAAATCATATATTCATTTCTGGCAATGATATCTCAACTTCACCTTATATTAATACAGC   | 180  |
| Db | 4802 | AGCAAAATCATATATTCATTTCTGGCAATGATATCTCAACTTCACCTTATATTAATACAGC   | 4861 |
| Qy | 181  | CATCATATTCATAGCTCGGCAACCAACCAAGTCACTAACAACCTGCAATCATACAAGA      | 240  |
| Db | 4862 | CATCATATTCATAGCTCGGCAACCAACCAAGTCACTAACAACCTGCAATCATACAAGA      | 4921 |
| Qy | 241  | TGCAAAAGCCAGATCAAGAACACAAACCCAAATATACATCAAAACCAACCATATAGCTTCAAC | 300  |
| Db | 4922 | TGCAAAAGCCAGATCAAGAACACAAACCCAAATATACATCAAAACCAACCATATAGCTTGG   | 4981 |
| Qy | 301  | AATCAGTCTTCCAACTGTCTGAAATTAATATCAACAAACCAACCATATAGCTTCAAC       | 360  |
| Db | 4982 | AATCAGTCTTCCAACTGTCTGAAATTAATATCAACAAACCAACCATATAGCTTCAAC       | 5041 |
| Qy | 361  | AACACGAGGTCAAGTCAAACTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT    | 420  |
| Db | 5042 | AACACGAGGTCAAGTCAAACTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT    | 5101 |
| Qy | 421  | AACCCAAACACAAACCCAGCAAGCCCTACTACAAACCAAGCCCAACCAACCAACCAACCA    | 480  |
| Db | 5102 | AACCCAAACACAAACCCAGCAAGCCCTACTACAAACCAAGCCCAACCAACCAACCAACCA    | 5161 |
| Qy | 481  | ACCCAAATATGATTTTCATTCGAAAGTGTAACTTTGTACCTGCAAGCATATGCAAGCA      | 540  |
| Db | 5162 | ACCCAAATATGATTTTCATTCGAAAGTGTAACTTTGTACCTGCAAGCATATGCAAGCA      | 5221 |
| Qy | 541  | CAATCCAACTGTGGGCTATCTGCAAAAGATACCAAAACCAAAACCAAGAAAGAAAC        | 600  |
| Db | 5222 | CAATCCAACTGTGGGCTATCTGCAAAAGATACCAAAACCAAAACCAAGAAAGAAAC        | 5281 |
| Qy | 601  | CACCACCAAGCCTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA         | 660  |
| Db | 5282 | CATTACCAAGCCCAAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA          | 5341 |
| Qy | 661  | AACCACTAAACCAAGGAAGTACCCACCAAGCCCAAGAGAGCCACCAATCAACAC          | 720  |
| Db | 5342 | AACCACTAAACCAAGGAAGTACCCACCAAGCCCAAGAGAGCCACCAATCAACAC          | 5401 |
| Qy | 721  | CACCAAAACAAACATCAACATCACTGCTCAACCAACCAAGGAATCCAAACT             | 780  |
| Db | 5402 | CACCAAAACAAACATCAACATCACTGCTCAACCAACCAAGGAATCCAAACT             | 5461 |

|    |      |   |      |
|----|------|---|------|
| Qy | 781  | CACAAGTCAAATGGAACCTTCCACTCAACTCCTCCGAGGCAATCTAAGCCCTTCTCA   | 840  |
| Db | 5462 | CACAAGTCAAATGGAACCTTCCACTCAACTCCTCCGAGGCAATCTAAGCCCTTCTCA   | 5521 |
| Qy | 841  | AGTCTCCACAACATCCGAGCACCCATCACAACCTCTATCTCCACCCCAACAACACGCA  | 900  |
| Db | 5522 | AGTCTCTACAACATCCGAGTACCCATCACAACCTTCTATCTCCACCCCAACAACACGCA | 5581 |
| Qy | 901  | GTAGTTATTAAAAAAA 919  |      |
| Db | 5582 | GTAGTTACTTAAAAACATA 5600                                    |      |

Search completed: October 30, 2003, 01:17:36  
Job time : 74.4618 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 16:50:52 ; Search time 313.982 Seconds  
(without alignments)  
7909.644 Million cell updates/sec

Title: US-09-462-816-1  
Perfect score: 920  
Sequence: 1 tgcacacatgtccaaaaca.....gtagttattataaaaaaaa 920

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                 |
|------------|-------|-------------|--------|----|-----------------------------|
| 1          | 920   | 100.0       | 920    | 14 | AAQ45686 Respiratory syncyt |
| 2          | 920   | 100.0       | 920    | 20 | AAQ08421 Membrane bound G p |
| 3          | 857.6 | 93.2        | 935    | 8  | AAV18736 Sequence encoding  |
| 4          | 857.6 | 93.2        | 935    | 19 | AAV18736 HRSV glycoprotein  |
| 5          | 856   | 93.0        | 935    | 13 | AAQ29623 HRSV glycoprotein  |
| 6          | 853.4 | 92.8        | 15222  | 18 | AAV18440 Human respiratory  |
| 7          | 853.4 | 92.8        | 15223  | 18 | AAV18440 Respiratory syncyt |
| 8          | 853.4 | 92.8        | 15223  | 19 | AAV17553 Respiratory syncyt |

|    |       |      |       |    |                               |
|----|-------|------|-------|----|-------------------------------|
| 9  | 853.4 | 92.8 | 15223 | 21 | AAV18440 Respiratory syncyt   |
| 10 | 848.2 | 92.2 | 918   | 22 | AAQ08421 Respiratory syncyt   |
| 11 | 837.8 | 91.1 | 897   | 24 | AAQ20145 Polynucleotide seq   |
| 12 | 830.4 | 90.3 | 15210 | 20 | AAQ08422 Respiratory syncyt   |
| 13 | 715   | 77.7 | 715   | 20 | AAQ08422 Respiratory syncyt   |
| 14 | 637.4 | 69.3 | 696   | 19 | AAV17552 Respiratory syncyt   |
| 15 | 441.4 | 48.0 | 15225 | 19 | AAV17552 Respiratory syncyt   |
| 16 | 441.4 | 48.0 | 15225 | 21 | AAV17552 Respiratory syncyt   |
| 17 | 441.4 | 48.0 | 15225 | 24 | AAV17552 Respiratory syncyt   |
| 18 | 422.2 | 45.9 | 15229 | 19 | AAV18276 Human respiratory    |
| 19 | 422.2 | 45.9 | 15229 | 20 | AAV18276 Human respiratory    |
| 20 | 422.2 | 45.9 | 15229 | 20 | AAV18276 Nucleotide sequence  |
| 21 | 411.4 | 44.7 | 15218 | 19 | AAV18276 DNA encoding the L   |
| 22 | 411.4 | 44.7 | 15218 | 20 | AAV18276 RSV isolate 2B wil   |
| 23 | 411.4 | 44.7 | 15218 | 20 | AAV18276 Nucleotide sequence  |
| 24 | 411.4 | 44.7 | 15218 | 20 | AAV18276 Nucleotide sequence  |
| 25 | 411.4 | 44.7 | 15219 | 19 | AAV18277 RSV vaccine 2B33F    |
| 26 | 411.4 | 44.7 | 15219 | 19 | AAV18277 RSV vaccine 2B20L    |
| 27 | 411.4 | 44.7 | 15219 | 19 | AAV18279 RSV revertant 2B33   |
| 28 | 411.4 | 44.7 | 15219 | 19 | AAV18280 RSV revertant 2B20   |
| 29 | 411.4 | 44.7 | 15219 | 20 | AAV18280 Nucleotide sequence  |
| 30 | 411.4 | 44.7 | 15219 | 20 | AAV18280 Nucleotide sequence  |
| 31 | 411.4 | 44.7 | 15219 | 20 | AAV18280 Nucleotide sequence  |
| 32 | 411.4 | 44.7 | 15219 | 20 | AAV18280 Nucleotide sequence  |
| 33 | 411.4 | 44.7 | 15219 | 20 | AAV18280 DNA encoding the L   |
| 34 | 411.4 | 44.7 | 15219 | 20 | AAV18280 DNA encoding the L   |
| 35 | 411.4 | 44.7 | 15219 | 20 | AAV18280 DNA encoding the L   |
| 36 | 212.8 | 23.1 | 1050  | 17 | AAV18280 Encodes Streptococ   |
| 37 | 212.8 | 23.1 | 1050  | 22 | AAV18280 Nucleotide sequence  |
| 38 | 212.8 | 23.1 | 1356  | 22 | AAV18280 RSV subgroup A clon  |
| 39 | 210.8 | 22.9 | 303   | 16 | AAV18280 RSV subgroup A prote |
| 40 | 210.8 | 22.9 | 303   | 17 | AAV18280 RSV sub-group A w    |
| 41 | 210.8 | 22.9 | 303   | 17 | AAV18280 Respiratory Syncyt   |
| 42 | 210.8 | 22.9 | 303   | 17 | AAV18280 RSV G protein anti   |
| 43 | 210.8 | 22.9 | 303   | 21 | AAV18280 DNA encoding a G2N   |
| 44 | 210.8 | 22.9 | 303   | 21 | AAV18280 Nucleotide sequence  |
| 45 | 210.8 | 22.9 | 303   | 22 | AAV18280 Nucleotide sequence  |

ALIGNMENTS

RESULT 1  
AAQ45686  
ID AAQ45686 standard; DNA; 920 BP.

|    |  |                                     |
|----|--|-------------------------------------|
| XX | AAQ45686;  |                                     |
| XX |  |                                     |
| XX | 25-MAR-2003 (updated)  |                                     |
| DT | 13-JAN-1994 (first entry)  |                                     |
| XX |  |                                     |
| DE | Respiratory syncytial virus (RSV) G gene.                          |                                     |
| DE |  |                                     |
| KW | PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine; |                                     |
| KW | ds.  |                                     |
| XX |  |                                     |
| OS | Respiratory syncytial virus.                                       |                                     |
| XX |  |                                     |
| PH | Key  | Location/Qualifiers                 |
| FT | CDS  | 8...901                             |
| FT |  | /*tag= a                            |
| FT |  | /product= RSV G protein             |
| FT | misc_feature   | 89..176                             |
| FT |  | /*tag= b                            |
| FT |  | /label= Transmembrane anchor domain |
| XX |  |                                     |
| PN | WO9314207-A1.  |                                     |
| PD | 22-JUL-1993.   |                                     |
| XX |  |                                     |
| PP | 05-JAN-1993; 93WO-CA00001.   |                                     |
| XX |  |                                     |
| PR | 06-JAN-1992; 92GB-0000117.   |                                     |

XX (CONN-) CONNAUGHT LAB LTD.  
 XX PI Ewasysbyn ME, Klein MH;  
 XX DR WPI; 1993-243222/30.  
 XX DR P-PSDB; AAR39286.  
 XX  
 PT Multimeric hybrid genes and their chimeric proteins - are  
 PT vaccines against multiple pathogenic infections e.g.  
 PT para-influenza virus and respiratory syncytial virus  
 PT  
 PS Claim 11; Figure 7A-7D; 80pp; English.  
 XX  
 CC A novel multimeric hybrid gene is used as a vaccine. The gene  
 CC consists of two gene sequences which are linked and encode antigenic  
 CC regions, these two sequences being derived from two different  
 CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus  
 CC (RSV)). The gene sequences that are particularly used are those  
 CC which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
 CC G proteins (AAQ45685, AAQ45686).  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;  
 Query Match 100.0%; Score 920; DB 14; Length 920;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e-211;  
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACATCTCCAAACCAAGCAGCAGCAGCGCTAGACACTAGAAAGACCTGGGA 60  
 Db 1 TGCACATCTCCAAACCAAGCAGCAGCAGCGCTAGACACTAGAAAGACCTGGGA 60  
 QY 61 CACTCTCAATCATTTATTATTATCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 120  
 Db 61 CACTCTCAATCATTTATTATTATCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 120  
 QY 121 AGCACAATACATATTCATTTCTGGCAATGATAATCTCAACTTCACTTATATTAATTCAGC 180  
 Db 121 AGCACAATACATATTCATTTCTGGCAATGATAATCTCAACTTCACTTATATTAATTCAGC 180  
 QY 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAAGCTGCAATCATACAGA 240  
 Db 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAAGCTGCAATCATACAGA 240  
 QY 241 TGCAACAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 300  
 Db 241 TGCAACAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 300  
 QY 301 AATCAGCTTCTCAATCTGTCTGAATTTATCATACAAACCAACCACCATAGCTTCAAC 360  
 Db 301 AATCAGCTTCTCAATCTGTCTGAATTTATCATACAAACCAACCACCATAGCTTCAAC 360  
 QY 361 AACACGAGGAGTCAAGTCAAACTCGCAACCAACCAAGTCACAGTCAAAACCAACAC 420  
 Db 361 AACACGAGGAGTCAAGTCAAACTCGCAACCAACCAAGTCACAGTCAAAACCAACAC 420  
 QY 421 AACCCAAACACACCCAGCAGCCACTACAAACCAACGCAAAACCAACCAACCA 480  
 Db 421 AACCCAAACACACCCAGCAGCCACTACAAACCAACGCAAAACCAACCAACCA 480  
 QY 481 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGTACCTGCGAGCATATGCGACAA 540  
 Db 481 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGTACCTGCGAGCATATGCGACAA 540  
 QY 541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAAAACCAAGGAAGAAC 600  
 Db 541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAAAACCAAGGAAGAAC 600  
 QY 601 CACCACCAAGCTTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 660  
 Db 601 CACCACCAAGCTTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 660

QY 661 AACCACTAAACCAAGGAAGTACCCACCAACCAAGGCCACAGAAAGGCCAACCATCAACAC 720  
 Db 661 AACCACTAAACCAAGGAAGTACCCACCAACCAAGGCCACAGAAAGGCCAACCATCAACAC 720  
 QY 721 CACCAAAACCAACATCACAACTACACTGCTCACCACCAACCAACCAAGGAATCCAAACT 780  
 Db 721 CACCAAAACCAACATCACAACTACACTGCTCACCACCAACCAACCAAGGAATCCAAACT 780  
 QY 781 CACAAGTCAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840  
 Db 781 CACAAGTCAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840  
 QY 841 AGTCTCCACAAATCGGAGCAGCCATCAACCTCATCTCCCAACCAACCAAGCCCA 900  
 Db 841 AGTCTCCACAAATCGGAGCAGCCATCAACCTCATCTCCCAACCAACCAAGCCCA 900  
 QY 901 GTAGTTATTAAAAA 920  
 Db 901 GTAGTTATTAAAAA 920

RESULT 2  
 AAX08421  
 ID AAX08421 standard; cDNA; 920 BP.  
 XX  
 AC AAX08421;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Membrane bound G protein gene of respiratory syncytial virus.  
 XX  
 KW G protein; respiratory syncytial virus; RSV; recombinant vector;  
 KW vaccine; immune response; immunogenicity; tPA; antibody;  
 KW tissue plasminogen activator; ss.  
 XX  
 OS Respiratory syncytial virus (RSV).  
 XX

Key Location/Qualifiers  
 CDS 8..904  
 FT /\*tag= a  
 FT /product= "Membrane bound G protein"  
 XX  
 PN WO9904010-A1.  
 PD 28-JAN-1999.  
 XX  
 PF 16-JUL-1998; 98WO-CA00697.  
 XX  
 PR 18-JUL-1997; 97US-0896442.  
 XX  
 PA (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PI Klein MH, Li X, Sambhara S;  
 XX  
 DR WPI; 1999-132254/11.  
 DR P-PSDB; AAW96313.  
 XX  
 CC Immunogenic composition for generating antibodies against  
 PT respiratory syncytial virus - comprises non-replicating vector  
 PT containing the protein G sequence, useful in protective vaccines and  
 PT to raise antibodies for diagnosis  
 XX  
 PS Claim 3; Figure 2; 67pp; English.  
 XX  
 CC The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G protein's  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when

CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods.

XX SQ Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;

Query Match 100.0%; Score 920; DB 20; Length 920;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-211;  
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGC AAA CAT GTG CCA AAAA CA AGG CCA AAG CCA CGC TAA GAC A C T A G A A A A G A C C T G G G A 60  
 Db 1 TGC AAA CAT GTG CCA AAAA CA AGG CCA AAG CCA CGC TAA GAC A C T A G A A A A G A C C T G G G A 60

Qy 61 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120  
 Db 61 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120

Qy 121 AGCAAAATCACATATCCATTTCTGGCAATGATATCTCAACTTCACTTATTAATTACAGC 180  
 Db 121 AGCAAAATCACATATCCATTTCTGGCAATGATATCTCAACTTCACTTATTAATTACAGC 180

Qy 181 CATCATATTATAGCTCGGCAACCAAGTCACTAACTCAACTGCAATCATACAAGA 240  
 Db 181 CATCATATTATAGCTCGGCAACCAAGTCACTAACTCAACTGCAATCATACAAGA 240

Qy 241 TGC AAC CAG CAG AT C A G A C A C A C C C A A C T A C T C A C T C A G A T T C T A G C T T G G 300  
 Db 241 TGC AAC CAG CAG AT C A G A C A C A C C C A A C T A C T C A C T C A G A T T C T A G C T T G G 300

Qy 301 AATCAGCTTCTCCAATCTGTCTGAATTTATACATCAAAACCAACCACTACTAGCTTCAAC 360  
 Db 301 AATCAGCTTCTCCAATCTGTCTGAATTTATACATCAAAACCAACCACTACTAGCTTCAAC 360

Qy 361 AACACAGGAGTCAAGTCAAACTGCAACCAACCACTACTAGCTTCAACCAAC 420  
 Db 361 AACACAGGAGTCAAGTCAAACTGCAACCAACCACTACTAGCTTCAACCAAC 420

Qy 421 AACCCAAACCAACCCAGCAGCCCACTACAAACAGCCCAACCAACCAACCAACAA 480  
 Db 421 AACCCAAACCAACCCAGCAGCCCACTACAAACAGCCCAACCAACCAACCAACAA 480

Qy 481 ACCCAATAATGATTTTCACTTTCGAAGTGTAACTTTGTACCTTGCAGCATATGAGCAA 540  
 Db 481 ACCCAATAATGATTTTCACTTTCGAAGTGTAACTTTGTACCTTGCAGCATATGAGCAA 540

Qy 541 CAATCCAACTCTGCTGGGCTATCTGCAAAAGATACCAAAACCAAGAAAGAAAC 600  
 Db 541 CAATCCAACTCTGCTGGGCTATCTGCAAAAGATACCAAAACCAAGAAAGAAAC 600

Qy 601 CACCACCAAGCTTACAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCTCA 660  
 Db 601 CACCACCAAGCTTACAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCTCA 660

Qy 661 AACCACTAAACCAAGAGTATCCCAACCAAGCCCAAGAGAGCCCAACCAAC 720  
 Db 661 AACCACTAAACCAAGAGTATCCCAACCAAGCCCAAGAGAGCCCAACCAAC 720

Qy 721 CACCAAAAACCAATCACTACATGCTCAACCAACCAACCAACCAACCAACCAACT 780  
 Db 721 CACCAAAAACCAATCACTACATGCTCAACCAACCAACCAACCAACCAACCAACT 780

Qy 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCGAAGGCAATTAAGCCCTTCTCA 840  
 Db 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCGAAGGCAATTAAGCCCTTCTCA 840

Qy 841 AGTCTCCAAACATCCGAGCACCCTTCAACCTCTATCTCCACCCCAACACAGCCCA 900  
 Db 841 AGTCTCCAAACATCCGAGCACCCTTCAACCTCTATCTCCACCCCAACACAGCCCA 900

Qy 901 GTAGTTATTAAAAA 920  
 |||||||

Db 901 GTAGTTATTAAAAA 920

RESULT 3

AAAT70784

ID AAAT70784 standard; cDNA; 935 BP.

XX AC AAAT70784;

XX DT 25-MAR-2003 (updated)

XX DT 05-APR-1991 (first entry)

XX DE Sequence encoding human respiratory syncytial virus (HRSV) A2 strain

XX DE G protein.

XX KV Vaccine; ss.

XX OS Human respiratory syncytial virus (HRSV).

XX FH Key Location/Qualifiers

XX FT CDS 16..913

XX FT /\*tag= a

XX FN WO8704185-A.

XX PD 16-JUL-1987.

XX PF 23-DEC-1986; 86WO-US02756.

XX PR 14-JAN-1986; 86US-0818740.

XX PA (UYN-) UNIV NORTH CAROLINA.

XX PA (WERTZ) WERTZ G W.

XX DR WPI; 1987-206300/29.

XX DR P-PSDB; AAP70845.

XX PT Vaccines for human respiratory virus - comprising proteins or

XX PT fragment encoded by a DNA sequence coding for human respiratory

XX PT syncytial virus proteins.

XX PS Disclosure; Chart 13; 57pp; English.

XX CC A novel plasmid which comprises a DNA sequence encoding this

XX CC protein, and the protein itself, are claimed, for use as HRSV

XX CC vaccines. The vaccine can be administered to pregnant women or to

XX CC women of child bearing age to stimulate maternal antibodies.

XX CC Infants can also be vaccinated at 2-3 months of age.

XX CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match 93.2%; Score 857.6; DB 8; Length 935;  
 Best Local Similarity 95.8%; Pred. No. 1.4e-196;  
 Matches 881; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 1 TGC AAA CAT GTG CCA AAAA CA AGG CCA AAG CCA CGC TAA GAC A C T A G A A A A G A C C T G G G A 60  
 Db 9 TGC AAA CAT GTG CCA AAAA CA AGG CCA AAG CCA CGC TAA GAC A C T A G A A A A G A C C T G G G A 68

Qy 61 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120  
 Db 69 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 128

Qy 121 AGCAAAATCACATATCCATTTCTGGCAATGATATCTCAACTTCACTTATTAATTACAGC 180  
 Db 129 AGCAAAATCACATATCCATTTCTGGCAATGATATCTCAACTTCACTTATTAATTACAGC 188

Qy 181 CATCATATTATAGCTCGGCAACCAAGTCACTAACTCACTTATTAATTACAGC 240  
 Db 189 CATCATATTATAGCTCGGCAACCAAGTCACTAACTCACTTATTAATTACAGC 248

Qy 241 TGC AAC CAG CAG AT C A G A C A C A C C C A A C T A C T C A C T C A G A T T C T A G C T T G G 300





```

Db      849 AGTCTCTACACATCGAGTACCCATCACAACTTCATCTCCACCCCAACACACACGCGCA 908
QY      901 GTAGTTATTAAAAA 920
Db      909 GTAGTTACTTAAAAA 928

RESULT 6
AAT78440
ID      AAT78440 standard; DNA; 15222 BP.
XX
AC      AAT78440;
XX
DT      17-MAR-1998 (first entry)
XX
DE      Human respiratory syncytial virus strain A2.
XX
KW      Antisense oligonucleotide; respiratory syncytial virus;
KW      RSV; treatment; infection; inhibition; strain A2; ss.
XX
OS      Human respiratory syncytial virus.
XX
PN      WO9729757-A1.
XX
PD      21-AUG-1997.
XX
PF      14-FEB-1997; 97WO-US02531.
XX
PR      15-FEB-1996; 96US-0011725.
XX
PA      (CLEV-) CLEVELAND CLINIC FOUND.
PA      (USSH) US NAT INST OF HEALTH.
XX
PI      Cirino NM, Li G, Silverman RH, Torrence PF, Xiao W;
XX      WPI; 1997-424748/39.
XX
PT      Polynucleotide containing sequence anti-sense to region of RSV -
PT      connected via a linker to an activator of RNaseL, used to treat RSV
PT      infections
XX
PS      Disclosure; Pages 47-51; 89pp; English.
XX
CC      The present sequence was used in the preparation of a novel
CC      polynucleotide, comprising an antisense oligonucleotide, with a
CC      hydroxy group at one end, that is complementary to 15-20 bases of
CC      the anti-genomic RNA strand of a respiratory syncytial virus (RSV),
CC      a linker attached to the OH-end of the antisense oligonucleotide and
CC      an oligonucleotide activator of RNaseL attached to the linker. The
CC      polynucleotide can be used to treat RSV infections, which can also
CC      be treated by administration of the antisense oligonucleotide, so as
CC      to form a complex with activated RNase L in vivo. The
CC      polynucleotide can be transported across the cell membranes without
CC      carriers or permeability agents, and once introduced destroys
CC      antisense target RNA. It also inhibits RSV infection in vitro in a
CC      superior manner to the conventional drug, ribavirin.
XX
SQ      Sequence 15222 BP; 5923 A; 2706 C; 2356 G; 4237 T; 0 other;

Query Match      92.8%; Score 853.4; DB 18; Length 15222;
Best Local Similarity 95.5%; Pred. No. 3.2e-195;
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY      1 TGCAAAACATGTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60
Db      4681 TGCAAAACATGTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 4740
QY      61 CACTCTCAATCATTTATTATTCATATCATCGGGCTTATATTAAGTTAAATCTTAATCTGT 120
Db      4741 CACTCTCAATCATTTATTATTCATATCATCGGGCTTATATTAAGTTAAATCTTAATCTGT 4800
QY      121 AGCACAAATCACATATTCATTCCTGGCAATGATATCTCAACTTCATATATATACAGC 180

```

```

Db      4801 AGCACAAATCACATTTCCATTTCTGGCAATGATATCTCAACTTCATATATATGCGAGC 4860
QY      181 CATCATATTTATAGCTCTGGCAAAACCAAAAGTCACTAAACATGCAATCATCAAGA 240
Db      4861 CATCATATTTATAGCTCTGGCAAAACCAAAAGTCACTAAACATGCAATCATCAAGA 4920
QY      241 TGCACACAGCCAGATCAAGACACAAACCCACATACCTCTCAGGATCTCAGCTTGG 300
Db      4921 TGCACACAGCCAGATCAAGACACAAACCCACATACCTCTCAGGATCTCAGCTTGG 4980
QY      301 AATCAGCTTTCTCCAATCTGTCTGAATTTACATCACAAACCCACATACCTCTCAGCTTCAAC 360
Db      4981 AATCAGTCCCTCTAATCCGTCTGAATTTACATCACAAACCCACATACCTCTCAGCTTCAAC 5040
QY      361 AACACAGGAGTCAAGTCAAACTTGCAACCCCAACAGTCAAGACTAAAGAAACCAACAAAC 420
Db      5041 AACACAGGAGTCAAGTCAAACTTGCAACCCCAACAGTCAAGACTAAAGAAACCAACAAAC 5100
QY      421 AACCCAAACACAAACCCAGCAAGCCCACTACAAACCAAGCCCAACAAACCAACCAACAA 480
Db      5101 AACTCAACACAAACCCAGCAAGCCCACTACAAACCAAGCCCAACAAACCAACCAACAA 5160
QY      481 ACCCAATATGATTTTCACTTCGAAAGTGTTTAACTTTGTACCTCTCAGCATATGCAAGAA 540
Db      5161 ACCCAATATGATTTTCACTTTGAAGTGTTCACCTTTGTACCTCTCAGCATATGCAAGAA 5220
QY      541 CAATCCAACTCTCTGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGGAAAGAAAC 600
Db      5221 CAATCCAACTCTCTGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGGAAAGAAAC 5280
QY      601 CACCAACAGCTTACAAACCAACCAACCTTCAAGCAACCAACCAACCAACCAACCAACCTCA 660
Db      5281 CACTACCAAGCCCAACCAACCAACCTTCAAGCAACCAACCAACCAACCAACCAACCTCA 5340
QY      661 AACCACTAAACCAAGGAGTACCCACCAACCAAGCCCAACCAAGGAGCCCAACCAACCAAC 720
Db      5341 AACCACTAAATCAAGGAGTACCCACCAACCAAGCCCAACCAAGGAGCCCAACCAACCAAC 5400
QY      721 CACCAAAACCAACATCAACATCTACTGCTCAACCAACCAACCAACCAAGGAAATCCAAACCT 780
Db      5401 CACCAAAACCAACATCAACATCTACTGCTCAACCAACCAACCAACCAAGGAAATCCCAAGACT 5460
QY      781 CACCACTCAATGGAACCACTTCCATCAACCTCTCCGAGGCAATCTTAAGCCCTTCTCA 840
Db      5461 CACAAGTCAAAATGGAACCACTTCCATCAACCTCTCCGAGGCAATCTTAAGCCCTTCTCA 5520
QY      841 AGTCTCCCAACATCCGAGCAGCCCATCAACCCCTCATCTCCACCCCAACCAACCAACGCGCA 900
Db      5521 AGTCTCTACACATCCGAGTACCCATCAACCTTCACTTCACTTCCGAGGCAATCTTAAGCCCTTCTCA 5580
QY      901 GTAGTTATTAAAAA 919
Db      5581 GTAGTTACTTAAAAACATA 5599

RESULT 7
AAT63430
ID      AAT63430 standard; DNA; 15223 BP.
XX
AC      AAT63430;
XX
DT      02-JUL-1997 (first entry)
XX
DE      Respiratory syncytial virus anti-genome.
XX
KW      RSV; vaccine; gene therapy; upper respiratory tract disease; ss.
XX
OS      Human respiratory syncytial virus strain A2.
XX
PN      WO9712032-A1.
XX
PD      03-APR-1997.

```

|          |          |  |      |
|----------|----------|--|------|
| Db       | 5222     | CAATCCAAACCTGCTGGGCTATCTCTCAAAAGATACCAAAACAAAAAACCCAGGAAAGAAAC       | 5282 |
| Qy       | 601      | CACCACCAAGCCTACAAAACCAACCAACCTTCAAGACGACCAAAAAGAGATCTCAAACTCA        | 660  |
| Db       | 5282     | CATTACCAAGGCCCAACAAAAAACAACCCCTCAAGACACCAAAAAAGATCCCAAACTCA          | 5341 |
| Qy       | 661      | AACCACTAAACCAAGGAAGTACCCACCAAGAGCCCAAGAGAGCCACCATCAACAC              | 720  |
| Db       | 5342     | AACCACTAAATCAAGGAAGTACCCACCAAGAGCCCAAGAGAGCCACCATCAACAC              | 5401 |
| Qy       | 721      | CACCAAAACCAACATCAACACTACATGTCTCAACAAACACACACAGGAAATCCAAAAC           | 780  |
| Db       | 5402     | CACCAAAACCAACATCATACTACTACTCTCACTCTCAACACCAAGGAATCCAGAAT             | 5461 |
| Qy       | 781      | CACAAGTCAAAATGGAAACCTTCCACTCAACTCTCTCGAAGGCAATTAAGCCCTTCTCA          | 840  |
| Db       | 5462     | CACAAGTCAAAATGGAAACCTTCCACTCAACTCTCTCGAAGGCAATTAAGCCCTTCTCA          | 5521 |
| Qy       | 841      | AGTCTCCACAACTACGAGCACCCTACCAAAACCTCATCTCCACCCCAACACACACGCCA          | 900  |
| Db       | 5522     | AGTCTCTCAACATCCGAGTACCCATCAACACTTCTCACTCAACCTTCACTCCACCCCAACACACGCCA | 5581 |
| Qy       | 901      | GTAGTTATTAACAAAAA 919  |      |
| Db       | 5582     | GTAGTTACTTAAAAACATA 5600   |      |
| RESULT 8 |          |  |      |
| AAV17553 |          |  |      |
| ID       | AAV17553 | standard; cdna; 15223 BP.  |      |
| XX       | XX       | AAV17553;  |      |
| XX       | XX       | 20-JUL-1998 (first entry)  |      |
| XX       | XX       | Respiratory syncytial virus antigenome.                              |      |
| XX       | XX       | RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.             |      |
| XX       | XX       | Human respiratory syncytial virus D46.                               |      |
| XX       | XX       | WO9802530-A1.  |      |
| XX       | XX       | 22-JAN-1998.   |      |
| XX       | XX       | 15-JUL-1997; 97WU-US12269.   |      |
| XX       | XX       | 23-MAY-1997; 97US-0047634.   |      |
| PR       | PR       | 15-JUL-1996; 96US-0021773.   |      |
| PR       | PR       | 09-MAY-1997; 97US-0046141.   |      |
| XX       | XX       | (USSH ) US DEPT HEALTH & HUMAN SERVICES.                             |      |
| PA       | PA       | Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN;               |      |
| PI       | PI       | Whitehead SS;  |      |
| XX       | XX       | WPI; 1998-110579/10.   |      |
| XX       | XX       | Attenuated respiratory syncytial virus vaccines - useful to protect  |      |
| PT       | PT       | individuals against RSV infection                                    |      |
| XX       | XX       | Example 7; Page 188-195; 238pp; English.                             |      |
| XX       | XX       | This is the 5'-3' positive sequence nucleotide sequence of           |      |
| CC       | CC       | respiratory syncytial virus (RSV) D46. The genome is                 |      |
| CC       | CC       | negative-sense; the complete nucleotide sequence of the              |      |
| CC       | CC       | wild-type B-1 virus has also been determined (see AAV17552).         |      |
| CC       | CC       | A novel infectious recombinant RSV comprises a RSV genome or         |      |
| CC       | CC       | antigenome, a major nucleocapsid (N) protein, a nucleocapsid         |      |
| CC       | CC       | phosphoprotein (P), a large polymerase protein (L), and a RNA        |      |
| CC       | CC       | polymerase elongation factor, where the recombinant RSV has at       |      |
| CC       | CC       | least two attenuating mutations, one of the mutations specifying a   |      |
| CC       | CC       | temperature-sensitive (ts) substitution at amino acid Dhe521         |      |



CC Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts  
 CC nucleotide substitution in the gene-start sequence of gene M2.  
 CC Also claimed are: (i) an isolated infectious RSV particle which  
 CC comprises a recombinant RSV (anti)genome, N, P, and L proteins, a  
 CC RNA polymerase elongation factor, where the (anti)genome is modified;  
 CC (i) to ablate or modulate expression of a SH, NS1, NS2 or G gene or  
 CC a cis-acting regulatory sequence; and (ii) by a termination codon  
 CC introduced within a selected gene, or by a change in sequence.  
 CC position or presence of a GS or GE transcription signal relative to  
 CC the selected gene; (2) an expression vector; and (3) an RSV strain  
 CC selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR  
 CC 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452),  
 CC cpts 530/1009 (ATCC VR 2451), or cpts 530/1030 (ATCC VR 2455), or  
 CC B-1 cp52/2B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated  
 CC attenuated recombinant RSV and RSV particles are used in a vaccine  
 CC to stimulate the immune system of an individual to induce  
 CC protection against RSV. The expression vector of (2) is used for  
 CC the production of infectious attenuated RSV particles.

XX SQ Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;

Query Match 92.8%; Score 853.4; DB 19; Length 15223;  
 Best Local Similarity 95.5%; Pred. No. 3.2e-195;  
 Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCAACATGTCGCAAAACCAAGGACCAAGCGACGACGCTAGACACTAGAAAGACCTGGGA 60  
 DB 4682 TGCAACATGTCGCAAAACCAAGGACCAAGCGACGACGCTAGACACTAGAAAGACCTGGGA 4741

QY 61 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT 120  
 DB 4742 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT 4801

QY 121 AGCAAAATCATATTCATTCATCTCTGGCAATGATATCTCACTTCACTTAATTTACAGC 180  
 DB 4802 AGCAAAATCATATTCATTCATCTCTGGCAATGATATCTCACTTCACTTAATTTACAGC 4861

QY 181 CATCATATTCATAGCTGGCAAAACCAAGGACCAAGGACCACTTAACCTGCAATCATACAAGA 240  
 DB 4862 CATCATATTCATAGCTGGCAAAACCAAGGACCAAGGACCACTTAACCTGCAATCATACAAGA 4921

QY 241 TGCAACAGCCAGATCAAGAACCAACCAACCCCAATACCTCTCAGGATCCTCAGCTTGG 300  
 DB 4922 TGCAACAGCCAGATCAAGAACCAACCAACCCCAATACCTCTCAGGATCCTCAGCTTGG 4981

QY 301 AATCAGTTCTCAATCTGCTCTGAAATTAATCATCAAAACCAACCACTAGCTTCAAC 360  
 DB 4982 AATCAGTTCTCAATCTGCTCTGAAATTAATCATCAAAATCAACCACTAGCTTCAAC 5041

QY 361 AACCCAGGAGTCAGTCAAACTGCAACCCCAACAGTCAAGACTTAAACCAACCAACCAAC 420  
 DB 5042 AACCCAGGAGTCAGTCAAACTGCAACCCCAACAGTCAAGACTTAAACCAACCAACCAAC 5101

QY 421 AACCCCAACCAACCAACCCAGGACCACTCAAAACCAACCCCAACCAACCAACCAACCA 480  
 DB 5102 AACTCAACCAACCAACCCAGGACCACTCAAAACCAACCCCAACCAACCAACCAACCA 5161

QY 481 ACCCAATATGATTTTCACTTGAAGTGTGTTAACTTTGACCTGCGAGATATGCGAGAA 540  
 DB 5162 ACCCAATATGATTTTCACTTGAAGTGTGTTAACTTTGACCTGCGAGATATGCGAGAA 5221

QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATATACCAAAACCAACCAACCAACCAACCA 600  
 DB 5222 CAATCCAACTGCTGGGCTATCTGCAAAAGATATACCAAAACCAACCAACCAACCAACCA 5281

QY 601 CACCAACCAAGCTTACAAACCAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 660  
 DB 5282 CACTACCAAGCCCAACCAAAACCAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 5341

QY 661 AACCACTTAAACCAAGGAAGTACCCCAACCAAGCCCAAGAGAGCCCAACCAATCAACAC 720  
 DB 5342 AACCACTTAAACCAAGGAAGTACCCCAACCAAGCCCAAGAGAGCCCAACCAATCAACAC 5401

QY 721 CACCAAAACAAACATCACACTACTACTGCTTCACTCAACCAACCAACAGGAATCCAAACT 780  
 DB 5402 CACCAAAACAAACATCATACTACTACTACTTCACTCACTCAACCAACCAACAGGAATCCAAACT 5461

QY 781 CACAAGTCAAAATGGAACCTTCCACTCACTCACTCACTCACTTCTCGAAGGCAATCTAAGCCCTTCTCA 840  
 DB 5462 CACAAGTCAAAATGGAACCTTCCACTCACTCACTCACTCACTTCTCGAAGGCAATCTCAAGCCCTTCTCA 5521

QY 841 AGTCTCCACAACTCCGAGCACCCATCACAAACCTCATCTCCACCCCAACCAACAGGCCA 900  
 DB 5522 AGTCTCTACAACTCCGAGTACCCATCAACAACTTCACTTCCACCCCAACCAACAGGCCA 5581

QY 901 GTAGTTATTAAAAAAA 919  
 DB 5582 GTAGTTACTTAAAAACATA 5600

RESULT 9  
 AAA88743  
 ID AAA88743 standard; cDNA; 15223 BP.  
 XX AC AAA88743;  
 XX DT 19-FEB-2001 (first entry)  
 XX DE Respiratory syncytial virus D46 5'-3' positive sense sequence.  
 XX KW RSV; vaccine; attenuation; pneumonia; bronchiolitis; mutant; ss.  
 XX OS Chimeric - Human respiratory syncytial virus.  
 XX OS Chimeric - Bacteriophage T7.  
 XX OS Synthetic.

EH Key Location/Qualifiers  
 FT variation replace(4,G)  
 FT mutation /\*tag= a  
 FT /\*tag= b  
 FT /\*note= "single C insertion to create AflII site  
 FT mutation replace(1139,A)  
 FT /\*tag= c  
 FT /\*note= "creates NcoI site in N gene nontranslated  
 FT mutation replace(1140,G)  
 FT /\*tag= d  
 FT /\*note= "creates NcoI site in N gene nontranslated  
 FT mutation replace(5612,A)  
 FT /\*tag= e  
 FT /\*note= "creates StuI site in G/F intergenic region"  
 FT mutation replace(5616,A)  
 FT /\*tag= f  
 FT /\*note= "creates StuI site in G/F intergenic region"  
 FT mutation replace(7560,A)  
 FT /\*tag= g  
 FT /\*note= "creates SphI site in F/M2 intergenic region"

XX WO200061611-A2.  
 PN 19-OCT-2000.  
 XX 31-MAR-2000; 2000WO-US08802.  
 XX 13-APR-1999; 99US-0291894.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Collins PL, Murphy BR, Whitehead SS;  
 XX WPI; 2000-679462/66.  
 XX Infectious chimeric respiratory syncytial virus (RSV) produced from  
 PT

PT cloned nucleotide sequences, useful as a vaccine against diseases  
 PT caused by the virus, such as pneumonia and bronchiolitis -

PS Example 7; Page 262-268; 280pp; English.

XX The present sequence is that of the 5' to 3' positive-sense  
 CC sequence of human respiratory syncytial virus (RSV); the genome  
 CC itself is negative-sense. This antigenome cDNA, termed D46, was  
 CC synthesized in segments by RT-PCR using synthetic oligonucleotides  
 CC as primers and intracellular RSV mRNA or genome RNA isolated from  
 CC purified viruses as template. The antigenome includes a 5'-terminal  
 CC nonviral G triplet contributed by the T7 promoter, 4 sequence  
 CC markers (see AAA88745-47) at positions 1099 (which adds 1 nucleotide  
 CC to the length), 1139, 5611 and 7559, a ribozyme and tandem T7  
 CC terminators, and a single nonviral 3'-phosphorylated residue  
 CC contributed to the 3' end by ribozyme cleavage. The invention  
 CC provides an isolated infectious chimeric RSV comprising a major  
 CC nucleocapsid (N) protein, a nucleocapsid phosphoprotein (P), a  
 CC large polymerase protein (L), an RNA polymerase elongation factor,  
 CC and a partial or complete RSV genome or antigenome of one RSV  
 CC strain or subgroup virus combined with a heterologous gene of a  
 CC different RSV strain or subgroup virus. The chimeric RSV is  
 CC infectious and attenuated, preferably by introduction of selected  
 CC mutations. It is useful as a vaccine against RSV, which causes  
 CC diseases such as pneumonia and bronchiolitis in infants. The  
 CC immune system of an individual is stimulated to induce protection  
 CC against natural RSV infection, preferably in a multivalent manner  
 CC to achieve protection against multiple RSV strains and/or subgroups.

XX SQ Sequence 15223 BP; 5921 A; 2705 C; 2361 G; 4236 T; 0 other;

Query Match 92.8%; Score 853.4; DB 21; Length 15223;  
 Best Local Similarity 95.5%; Pred. No. 3.2e-195;  
 Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCACATCTCCAAAACAAAGACACAAACGACGCTTAAGACACTAGAAAAGACCTGGGA 60  
 DB 4682 TGCACATCTCCAAAACAAAGACACAAACGACGCTTAAGACACTAGAAAAGACCTGGGA 4741  
 QY 61 CACTCTCAATCAATTAATTAATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 120  
 DB 4742 CACTCTCAATCAATTAATTAATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 4801  
 QY 121 AGCACAATCAATTAATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 180  
 DB 4802 AGCACAATCAATTAATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 4861  
 QY 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAACTGCAATCATACAAGA 240  
 DB 4862 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAACTGCAATCATACAAGA 4921  
 QY 241 TGCAACAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 300  
 DB 4922 TGCAACAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 4981  
 QY 301 AATCAGTCTTCAATCTGTCTGAATTAATCATCAACCAACCACTACTAGCTTCAAC 360  
 DB 4982 AATCAGTCTTCAATCTGTCTGAATTAATCATCAACCAACCACTACTAGCTTCAAC 5041  
 QY 361 AACACCAAGGATCAAGTCAAACTGCAACCCCAACCAAGTCACACTAACTGCAATCAAC 420  
 DB 5042 AACACCAAGGATCAAGTCAAACTGCAACCCCAACCAAGTCACACTAACTGCAATCAAC 5101  
 QY 421 AACCCAAACCAACCCAGCAAGCCCACTACAAAACCAACCGCAAAACCAACCAACCA 480  
 DB 5102 AACTCAACCAACCAACCCAGCAAGCCCACTACAAAACCAACCGCAAAACCAACCAACCA 5161  
 QY 481 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGACCTCGAGCATATGCGCAA 540  
 DB 5162 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGACCTCGAGCATATGCGCAA 5221  
 QY 541 CAAATCAACCTGCTGGGCTATCTGCAAAAGAAATACCAACCAAAACCCAGGAAGAAC 600

DB 5222 CAAATCAACCTGCTGGGCTATCTGCAAAAGAAATACCAACCAAAACCCAGGAAGAAAC 5281  
 QY 601 CACCAACCAAGCTTACAAAAAACCAACCTTCAAGACAAACCAAAAAAGATCTTCAAACTCA 660  
 DB 5282 CACTACCAAGCCCAACAAAAAACCAACCTTCAAGACAAACCAAAAAAGATCTTCAAACTCA 5341  
 QY 661 AACCACTAAACCAAGGAAGTACCCACCAAGCCCAAGAGCCCAACCAATCAACAC 720  
 DB 5342 AACCACTAAATCAAGGAAGTACCCACCAAGCCCAAGAGCCCAACCAATCAACAC 5401  
 QY 721 CACCAAAACCAACATCACAACTACATCTGCTCACCAACCAACCAAGGAAGTCCAAACT 780  
 DB 5402 CACCAAAACCAACATCACAACTACATCTGCTCACCAACCAACCAAGGAAGTCCAAACT 5461  
 QY 781 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTTAAGCCCTTCTCA 840  
 DB 5462 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTTAAGCCCTTCTCA 5521  
 QY 841 AGTCTCCCAACATCCGAGCAGCCCATCAACCTCTATCTCCACCAACCAACAGCCCA 900  
 DB 5522 AGTCTCTCAACATCCGAGTACCCATCAACCTCTATCTCCACCAACCAACAGCCCA 5581  
 QY 901 GTAGTTATTAAAAA 919  
 DB 5582 GTAGTTATTAAAAACATA 5600  
 RESULT 10  
 AAC88494  
 ID AAC88494 standard; DNA; 918 BP.  
 AC AAC88494;  
 XX DT 12-MAR-2001 (first entry)  
 XX Human RSV G-protein gene.  
 DE Chimeric; respiratory syncytial virus; RSV; immunize; ds.  
 KW Unidentified.  
 OS WO200068392-A1.  
 XX PD 16-NOV-2000.  
 XX PF 10-MAY-2000; 2000WO-US12582.  
 XX PR 11-MAY-1999; 99US-0133536.  
 XX PA (UNII ) UNIV ILLINOIS FOUND.  
 XX PI Buetow DE, Korban SS, Sandhu J, Krasnyanski SF;  
 XX WPI; 2001-122707/13.  
 XX Chimeric nucleic acid construct for immunizing animals and humans  
 PT against respiratory syncytial virus (RSV), comprises a sequence adapted  
 PT for expression in plants and a RSV protein or peptide coding sequence  
 PT -  
 XX Disclosure; Fig 10; 67pp; English.  
 PS The present invention relates to a chimeric nucleic acid construct  
 CC comprising: a nucleotide sequence adapted for protein expression in  
 CC plants; and a respiratory syncytial virus (RSV) coding sequence  
 CC encoding an RSV protein or an antigenic protein or peptide of RSV.  
 CC The construct can be used to immunize animals and humans against  
 CC respiratory syncytial virus. The use of transgenic plants to  
 CC generate the antigen allows the production of greater amounts of  
 CC antigen.  
 XX Sequence 918 BP; 366 A; 293 C; 100 G; 159 T; 0 other;

Query Match 92.2%; Score 848.2; DB 22; Length 918;  
Best Local Similarity 95.8%; Pred. No. 2.6e-194;  
Matches 871; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```
QY 1 TGCACACATGTCACAAAACAGAGGACCAAGCCGCTAGAGACACTAGAAAGACCTGGGA 60
DB 9 TGCACACATGTCACAAAACAGAGGACCAAGCCGCTAGAGACACTAGAAAGACCTGGGA 68
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGCTTTATATAGTTAAATCTTAAATCTGT 120
DB 69 CACTCTCAATCATTTATTTATTTATTCATATCATCGGCTTTATATAGTTAAATCTTAAATCTGT 128
QY 121 AGCACAATCAGATTTATTCATTTCTGCGAATGATAATCTCAACTTCTATTATATACAGC 180
DB 129 AGCACAATCAGATTTATTCATTTCTGCGAATGATAATCTCAACTTCTATTATATTCAGC 188
QY 181 CATCATATTTCATAGCTTCGGCAACCAACCAAGTACACTAACTGCAATCATACAAGA 240
DB 189 CATCATATTTCATAGCTTCGGCAACCAACCAAGTACACTAACTGCAATCATACAAGA 248
QY 241 TGCACAAGCCAGATCAAGAACAACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 300
DB 249 TGCACAAGCCAGATCAAGAACAACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 308
QY 301 ATTCAGCTTCTCAATCTGCTGAAATTCATACCAACCAACCACTACCTAGCTTCAAC 360
DB 309 ATTCAGCTTCTCAATCTGCTGAAATTCATACCAACCAACCACTACCTAGCTTCAAC 368
QY 361 AACACACAGGAGTCAAGTCAAACTGCAACCCCAACCAAGTCAAGACTAAAAACAACAAC 420
DB 369 AACACACAGGAGTCAAGTCAAACTGCAACCCCAACCAAGTCAAGACTAAAAACAACAAC 428
QY 421 AACCCAAACAACCCAGGAGCCACTACAAACAAAGCCCAACCAACCAACCAACCAAC 480
DB 429 AACCCAAACAACCCAGGAGCCACTACAAACAAAGCCCAACCAACCAACCAACCAAC 488
QY 481 ACCCAATATGATTTTCACTTCGAGTGTGTTAACTTTGATCCCTGAGCATATGCAGCA 540
DB 489 ACCCAATATGATTTTCACTTCGAGTGTGTTAACTTTGATCCCTGAGCATATGCAGCA 548
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATATACCAACCAAAACCCAGGAAGAAAC 600
DB 549 CAATCCAACTGCTGGGCTATCTGCAAAAGATATACCAACCAAAACCCAGGAAGAAAC 608
QY 601 CACCAACCAAGCTTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 660
DB 609 CACTTACCAAGCCACAAAACCAACCCCTCAAGACCAACCAAAAGATCCCAAACTCA 668
QY 661 AACCACTAAACCAAGGAGTACCCACCAACCAAGCCCAAGAGGACCAACCACTCAAC 720
DB 669 AACCACTAAACCAAGGAGTACCCACCAACCAAGCCCAAGAGGACCAACCACTCAAC 728
QY 721 CACCAAAACCAACATCAACACTACACTGCTCACCACCAACCAACCAAGGAATCCAAACT 780
DB 729 CACCAAAACCAACATCAACACTACACTGCTCACCACCAACCAACCAAGGAATCCAGACT 788
QY 781 CACAAGTCAATGGAACCTTCCACTCAACTCTCTCCGAAGGCAATTAAGCCCTTCTCA 840
DB 789 CACAAGTCAATGGAACCTTCCACTCAACTCTCTCCGAAGGCAATTAAGCCCTTCTCA 848
QY 841 AGTCTCCACACATCCGAGGACCCATCAACCTCTGATCTTCCACCAACCAACCAAGCCA 900
DB 849 AGTCTCTACACATCCGAGGACCCATCAACCTCTGATCTTCCACCAACCAACCAAGCCA 908
QY 901 GTAGTTATT 909
DB 909 GTAGTTACT 917
```

RESULT 11  
AAS20145  
ID AAS20145 standard; DNA; 897 BP.

```
AC AAS20145;
XX 09-APR-2002 (first entry)
XX Respiratory syncytial virus G protein DNA.
XX RSV; ds; G protein; heavily glycosylated protein; antianseimic; antiviral;
XX vaccine; gene therapy; paramyxovirus; sendai virus; PMV;
XX antiviral chemotherapeutic compound; humoral response;
XX cellular immune response; hPIV; paediatric respiratory disease;
XX globin gene transfer; sickle cell disease; beta-thalassaemia;
XX human immunodeficiency virus infection; HIV.
XX Human respiratory syncytial virus.
XX Key Location/Qualifiers
XX CDS 1..897
XX /*tag= a
XX /product= "G protein"
XX WO200192548-A2.
XX 06-DEC-2001.
XX 22-MAY-2001; 2001WO-US16610.
XX 01-JUN-2000; 2000US-208701P.
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX Portner A, Takimoto T;
XX WPI; 2002-130534/17.
XX P-PSDB; AAU74676.
XX Recombinant Sendai virus useful in vaccines to protect infection by
XX paramyxoviruses, comprises exogenous nucleic acid encoding
XX paramyxovirus protein or its antigenic fragment -
XX Disclosure; Page 47; 57pp; English.
XX The invention relates to a recombinant Sendai virus comprising an
XX exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its
XX antigenic fragment. The virus may be administered in combination
XX with an antiviral chemotherapeutic compound. Two or more viruses
XX expressing different PMV proteins may be co-administered. Compositions
XX comprising the virus are useful for eliciting a humoral and/or
XX cellular immune response to a PMV in a mammal, particularly a human.
XX Further a recombinant Sendai virus comprising an exogenous nucleic acid
XX encoding a second PMV protein is also administered and priming and/or
XX boosting humoral or cellular immune response comprises administering
XX one or more of a recombinant or isolated PMV protein or its antigenic
XX fragment, a DNA vaccine encoding the same, and a non-Sendai viral
XX vector encoding a PMV protein. The recombinant virus is useful as an
XX effective vaccine against hPIV or RSV (the major causes of paediatric
XX respiratory disease) and also to express any gene of
XX interest in target cells, providing a positive medical impact on
XX impaired cells. Wild-type globin gene transfer (i.e. gene therapy)
XX into stem cells effects a cure for sickle cell disease or beta-
XX thalassaemia. The recombinant virus may also prove effective in
XX conferring immunity to human immunodeficiency virus (HIV) infection.
XX The Sendai virus replicates at level that is high enough to
XX induce sufficient immunity, but does not cause any harm to human
XX recipient. The present sequence encodes a respiratory syncytial
XX virus (RSV) G protein (heavily glycosylated protein), a PMV protein
XX suitable for expression by the recombinant virus of the invention.
XX Sequence 897 BP; 359 A; 289 C; 95 G; 154 T; 0 other;
```

Query Match 91.1%; Score 837.8; DB 24; Length 897;  
Best Local Similarity 95.9%; Pred. No. 8e-192;  
Matches 860; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
QY 8 ATGTCCAAAAAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGACACTCTC 67
Db |||||
QY 1 ATGTCCAAAAAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGACACTCTC 60
Db |||||
QY 68 AATCAATTATTTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGTAGCAAA 127
Db |||||
QY 61 AATCAATTATTTATTCATATCATCGGTCTTATATAAGTTAAATCTTAAATCTGTAGCAAA 120
Db |||||
QY 128 ATCACTATTCCTTCCTGGCAATGATATCTCACTTCATCTTATATTAATTAAGCCATCATTA 187
Db |||||
QY 121 ATCACTATTCCTTCCTGGCAATGATATCTCACTTCATCTTATATTAATTAAGCCATCATTA 180
Db |||||
QY 188 TTCAATAGCTCGGCAACCAACCAAGTCACTAATCACTGCAATCATACAGATGCAACA 247
Db |||||
QY 181 TTCAATAGCTCGGCAACCAACCAAGTCACTAATCACTGCAATCATACAGATGCAACA 240
Db |||||
QY 248 AGCAGATCAAGAACCAACCCCAATACATCTCTCAGGATCTCTCAGCTTGGAAATCAGC 307
Db |||||
QY 241 AGCAGATCAAGAACCAACCCCAATACATCTCTCAGGATCTCTCAGCTTGGAAATCAGT 300
Db |||||
QY 308 TTCTCCAACTGTCTGAAATTCATACACACCAACCACTACTAGCTTCAACCAACCA 367
Db |||||
QY 301 CCTCTTAATCCGTCTGAAATTCATACCAATCACTACCACTACTAGCTTCAACCAACCA 360
Db |||||
QY 368 GGAATCAAGTCAAACTCTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 427
Db |||||
QY 361 GGAATCAAGTCAAACTCTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
Db |||||
QY 428 ACACCAACCAAGGACCACTCAAAACCAAGGCAACCAACCAACCAACCAACCAACCAAT 487
Db |||||
QY 421 ACACCAACCAAGGACCACTCAAAACCAAGGCAACCAACCAACCAACCAACCAACCAAT 480
Db |||||
QY 488 AATGATTTTCACTTCGAAGTGTATTAATTTGATCTCTGAGCATATGCGACGACATCA 547
Db |||||
QY 481 AATGATTTTCACTTCGAAGTGTATTAATTTGATCTCTGAGCATATGCGACGACATCA 540
Db |||||
QY 548 ACTGTCTGGGTATCTGCAAAAGAAATCAACCAACCAACCAAGGAAAGAAACCAACCA 607
Db |||||
QY 541 ACTGTCTGGGTATCTGCAAAAGAAATCAACCAACCAACCAAGGAAAGAAACCAACCA 600
Db |||||
QY 608 AAGCTTACAAAAAACCCTTCAGAACCAACCAACCAACCAACCAACCAACCAACCACT 667
Db |||||
QY 601 AAGCCACAAAAAACCCTTCAGAACCAACCAACCAACCAACCAACCAACCAACCACT 660
Db |||||
QY 668 AAACCAAGGAGTACCCACCAACCAAGGCAACCAAGGAGGCAACCAACCAACCAACCA 727
Db |||||
QY 661 AATCAAGGAGTACCCACCAACCAAGGCAACCAAGGAGGCAACCAACCAACCAACCA 720
Db |||||
QY 728 ACAAACATCAAACTACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAAGT 787
Db |||||
QY 721 ACAAACATCAAACTACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAAGT 780
Db |||||
QY 788 CAAATGGAAACCTTCGACCTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 847
Db |||||
QY 781 CAAATGGAAACCTTCGACCTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCT 840
Db |||||
QY 848 ACAACATCGAGACCCCATCAACCAACCTCACTCTCCACCAACCAACCAACCAACCAAGT 904
Db |||||
QY 841 ACAACATCGAGTACCCATCAACCTCACTCTCCACCAACCAACCAACCAACCAAGT 897
Db |||||
```

RESULT 12

AA59703/c

ID AA59703 standard; DNA; 15210 BP.

XX AA59703;

AC AA59703;

XX 26-JUL-1999 (first entry)

XX Polynucleotide sequence of RSV strain A2.

DE Antisense oligonucleotide; negative-strand RNA virus; activator; RNase L;

XX respiratory syncytial virus; RSV; influenza; mumps; rabies; ss.

KW

KW

KW

KW

KW

KW

KW

KW

KW

XX  
OS  
XX  
PN  
XX  
PD  
XX  
XX  
PF  
XX  
PR  
XX  
PA  
XX  
PA  
XX  
PI  
XX  
PI  
XX  
DR  
XX  
PT  
XX  
PS  
XX  
CC  
XX  
CC  
XX  
CC  
XX  
CC  
XX  
CC  
XX  
CC  
XX  
CC  
XX  
CC  
XX  
SQ

Query Match

Best Local Similarity 90.3%; Score 830.4; DB 20; Length 15210;

Matches 869; Conservative 3; Mismatches 41; Indels 6; Gaps 1;

QY 1 TGCACAAATGTCCTCAAAACCAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60  
Db 10530 TGCACAAATGTCCTCAAAACCAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 10471  
QY 61 CACTCTCAATCATTTTATTTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT 120  
Db 10470 CACTCTCAATCATTTTATTTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT 10417  
QY 121 AGCACAATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 180  
Db 10416 AGCACAATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 10357  
QY 181 CATCATATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 240  
Db 10356 CATCATATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 10297  
QY 241 TGCACAAAGCCAGATCAAGAACCAACCCCAACCAACCAACCAACCAACCAACCAACCAAGT 300  
Db 10296 TGCACAAAGCCAGATCAAGAACCAACCCCAACCAACCAACCAACCAACCAACCAACCAAGT 10237  
QY 301 AATCAGTCTTCTTAATTCGCTCTGAAATTTACATCAACCAACCAACCAACCAACCAACCAAGT 360  
Db 10236 AATCAGTCTTCTTAATTCGCTCTGAAATTTACATCAACCAACCAACCAACCAACCAAGT 10177  
QY 361 AACACCAAGGATCAAGTCAAACTTGCACCAACCAACCAACCAACCAACCAACCAACCAAGT 420  
Db 10176 AACACCAAGGATCAAGTCAAACTTGCACCAACCAACCAACCAACCAACCAACCAAGT 10117  
QY 421 AACCCAAACCAACCCAGAGCCCACTTACAAAAACCAACCAACCAACCAACCAACCAAGT 480  
Db 10116 AACCCAAACCAACCCAGAGCCCACTTACAAAAACCAACCAACCAACCAACCAAGT 10057

Respiratory syncytial virus.

WO922742-A1.

14-MAY-1999.

02-NOV-1998; 98WO-US23391.

03-NOV-1997; 97US-0962690.

(CLEV-) CLEVELAND CLINIC FOUND.

(USSH) US NAT INST OF HEALTH.

Cirino NM, Li G, Player MR, Silverman RH, Torrence PF;

Xiao W;

WPI; 1999-326917/27.

New composition useful for inhibiting or treating infections against negative-strand RNA virus

Disclosure; Fig 1; 98pp; English.

The specification describes a composition comprising a polynucleotide consisting of an antisense oligonucleotide containing a hydroxy group, complementary to the genomic or antigenomic strand of a negative-strand RNA virus; and an activator of RNase L. The polynucleotide is used to inhibit, or treat, infection by negative-strand RNA viruses, specifically respiratory syncytial virus (RSV) but also (para)influenza, mumps, and rabies. The polynucleotide can cross cell membranes without requiring carriers or permeabilizing agents, and can selectively cleave the RNA targeted by the oligonucleotide. The present sequence represents the polynucleotide sequence of RSV strain A2.

Sequence 15210 BP; 4232 A; 2351 C; 2700 G; 5919 T; 8 other;

QY 481 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGTTACCTGCGAGCATATGCGACAA 540  
 Db 10056 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGTTACCTGCGAGCATATGCGACAA 9997  
 QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGAAAGAAAC 600  
 Db 9996 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAGAAAGAAAC 9937  
 QY 601 CACACCAAGCTTACAAACCAACCTTCAAGACCAACCAAAAGAAATCTCAAACTCA 660  
 Db 9936 CACTTACCAAGCTTACAAACCAACCTTCAAGACCAACCAAAAGAAATCTCAAACTCA 9877  
 QY 661 AACCACTTAAACCAAGCTTACCAACCAAGCTTCAAGACCAACCAAAAGAAATCTCAAC 720  
 Db 9876 AACCACTTAAACCAAGCTTACCAACCAAGCTTCAAGACCAACCAAAAGAAATCTCAAC 9817  
 QY 721 CACCAAAACCAACCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 780  
 Db 9816 CACCAAAACCAACCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 9757  
 QY 781 CACAAGTCAATGAAACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 840  
 Db 9756 CACAAGTCAATGAAACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 9697  
 QY 841 AGTCTCCACACATCCGAGCACCCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 900  
 Db 9696 AGTCTCTACACATCCGAGTACCCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 9637  
 QY 901 GTAGTTATTAAAAAAA 919  
 Db 9636 GTAGTTATTAAAAACATA 9618

## RESULT 13

AAK08422

ID AAK08422 standard; cDNA; 715 BP.

XX

AC AAK08422;

XX

DT 28-JUN-1999 (first entry)

XX G protein gene fragment of respiratory syncytial virus.

XX G protein: respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;

XW tissue plasminogen activator; ss.

XX

OS Respiratory syncytial virus (RSV).

XX

PH Location/Qualifiers

FT CDS

FT 1..702

FT /\*tag= a

FT /product= "Secreted G protein"

XX

PN WO9904010-A1.

XX

PD 28-JAN-1999.

XX

PF 16-JUL-1998; 98WO-CA00697.

XX

PR 18-JUL-1997; 97US-0896442.

XX

PA (CONN-) CONNAUGHT LAB LTD.

XX

PI Klein MH, Li X, Sambhara S;

XX

DR WPI; 1999-132254/11.

DR

DR P-PSDB; AAW96314.

XX

XX Immunogenic composition for generating antibodies against

PT respiratory syncytial virus - comprises non-replicating vector

PT containing the protein G sequence, useful in protective vaccines and

PT to raise antibodies for diagnosis

PT

XX Claim 8; Figure 3; 67pp; English.  
 PS

XX The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods.

XX Sequence 715 BP; 305 A; 240 C; 70 G; 100 T; 0 other;

Query Match 77.7%; Score 715; DB 20; Length 715;

Best Local Similarity 100.0%; Pred. No. 2.5e-162;

Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 CACAAAGTCACACTCACTCACTCAGGATCTCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 265  
 Db 1 CACAAAGTCACACTCACTCACTCAGGATCTCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 60  
 QY 266 ACCCCAACTATCTCTCAGGATCTCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 325  
 Db 61 ACCCCAACTATCTCTCAGGATCTCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 120  
 QY 326 ATTACATCAAAACCAACCACTATCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 385  
 Db 121 ATTACATCAAAACCAACCACTATCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 180  
 QY 386 CACCCCACTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 445  
 Db 181 CACCCCACTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240  
 QY 446 ACTACAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
 Db 241 ACTACAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 QY 506 GTGTTTAACTTTGTACCTTGCAGCATATGCGAGCAATCCAACTTCAAGCTGTGGGTATCTGC 565  
 Db 301 GTGTTTAACTTTGTACCTTGCAGCATATGCGAGCAATCCAACTTCAAGCTGTGGGTATCTGC 360  
 QY 566 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
 Db 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 QY 626 ACCTTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
 Db 421 ACCTTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
 QY 686 ACCACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 745  
 Db 481 ACCACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
 QY 746 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805  
 Db 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
 QY 806 TCAACCTCTCCGAGGCAATCTTAAGCCCTTCTCAAGTCTCCACACATCCGAGCAACCAACCAACCAACCA 865  
 Db 601 TCAACCTCTCCGAGGCAATCTTAAGCCCTTCTCAAGTCTCCACACATCCGAGCAACCAACCAACCAACCA 660  
 QY 866 TCAACCTCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920  
 Db 661 TCAACCTCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715

RESULT 14

AAV38298  
 ID AAV38298 standard; DNA; 696 BP.  
 AC AAV38298;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE Respiratory syncytial virus glycoprotein G gene portion.  
 XX  
 XX RSV; glycoprotein G; pneumococcal surface protein A; PspA;  
 KW infection; Streptococcus pneumoniae; sepsis; otitis media;  
 KW meningitis; bacteraemia; pneumonia; vaccine; genetic immunisation;  
 KW ss.  
 XX  
 XX Respiratory syncytial virus.  
 OS  
 XX WO9824927-A1.  
 PN  
 XX 11-JUN-1998.  
 PD  
 XX 04-DEC-1997; 97WO-US22847.  
 PF  
 XX 04-DEC-1996; 96US-0759505.  
 PR  
 XX (UVAL-) UNIV ALABAMA.  
 PA  
 XX Briles DE, Curriel DT, McDaniel LS;  
 XX  
 XX WPI; 1998-333343/29.  
 DR  
 XX Plasmid containing pneumococcal epitope for expression in eukaryotic  
 XX cells - useful for eliciting immunological response to pneumococcal  
 PT infection or sepsis  
 PT  
 XX Example 1; Fig 1C; 47pp; English.  
 PS  
 XX This is a portion of the respiratory syncytial virus glycoprotein G  
 CC (RSVG) gene. It has been inserted into plasmid pCDNA3 (see AAV38297)  
 CC to create plasmid pGT4. This plasmid contains a human  
 CC cytomegalovirus immediate early promoter and the RSVG gene portion  
 CC such that when an in-frame fusion is made, the resultant fusion  
 CC protein may be transported to, and anchored in, a mammalian cell  
 CC membrane where it can be exposed to the host immune system.  
 CC Insertion of pneumococcal surface protein A (PspA) coding sequence  
 CC created plasmid pKSD2601. Intramuscular immunisation of BALB/c  
 CC mice with pKSD2601 induced protection against an otherwise lethal  
 CC challenge with a capsular type 3 pneumococcus. A claimed plasmid  
 CC for expression of pneumococcal epitope DNA in eukaryotic cells  
 CC includes a promoter for driving expression in a eukaryotic cell  
 CC (e.g. HCMV-IE). DNA encoding a leader sequence (e.g. of RSVG)  
 CC and DNA encoding a pneumococcal epitope such as PspA. The invention  
 CC also provides a vaccine comprising the plasmid and a suitable  
 CC carrier or diluent, and optionally one or more cytokines or DNA  
 CC encoding them, or a bacterial delivery system. The vaccine is used  
 CC to elicit an immunological response in a host, including humans,  
 CC susceptible to pneumococcal infection or sepsis. The plasmid can  
 CC also be used to express a pneumococcal epitope of interest in vitro.  
 XX  
 SQ Sequence 696 BP; 284 A; 211 C; 78 G; 123 T; 0 other;  
 Query Match 69.3%; Score 637.4; DB 19; Length 696;  
 Best Local Similarity 96.1%; Pred. No. 1.1e-143;  
 Matches 664; Conservative 0; Mismatches 26; Indels 1; Gaps 1;  
 8 ATGTCGAAACCAAGGACCAACCGCGTACGACACTAGAAAGACCTGGGACACTCTC 67  
 1 ATGTCGAAACCAAGGACCAACCGCGTACGACACTAGAAAGACCTGGGACACTCTC 60  
 68 AATCATTATTATTCATATCATCGGCTTATATAGTTAACTTAAATCTGTAGACAA 127  
 61 AATCATTATTATTCATATCATCGGCTTATATAGTTAACTTAAATCTGTAGACAA 120  
 128 ATCACCATTATCCATTCTGGCAATGATAATCTCACTTATATATAGCCATCAT 187

Db 121 ATCACCATTATCCATTCTGGCAATGATAATCTCACTTATATATAGCCATCAT 180  
 Qy 188 TTCATAGCCTCGGCAACCAACCAAGTCACTAACTGCAATCATCAAGATGCAACA 247  
 Db 181 TTCATAGCCTCGGCAACCAACCAAGTCACTAACTGCAATCATCAAGATGCAACA 240  
 Qy 248 AGCCAGATCAAGAACACACCCCAATACCTCACTCAGGATCTCAGTTGGAATCAGC 307  
 Db 241 AGCCAGATCAAGAACACACCCCAATACCTCACCAGAACTCTCAGTTGGAATCAGT 300  
 Qy 308 TTCTCCAAATCTGCTGAAATTCATATCAACCAACCACTACTAGTTTCAACACCA 367  
 Db 301 CCCTCTAATCCGTCGTAATTCATCAAAATCAACCACTACTAGTTTCAACACCA 360  
 Qy 368 GGAGTCAAGTCAAACTCTGCAACCCCAACCACTCAAGTCAAGTCAAAACCAACCAAC 426  
 Db 361 GGAGTCAAGTCAAACTCTGCAATCCACCACTCAAGTCAAGTCAAAACCAACCAAC 420  
 Qy 427 AACACACCCAGCAGCCCACTCAAAACCAACCAACCAACCAACCAACCAACCAAC 486  
 Db 421 AACACACCCAGCAGCCCACTCAAAACCAACCAACCAACCAACCAACCAACCAAC 480  
 Qy 487 TAATGATTTTCACTTCGAACTGTTTAACTTTGTAACCTGCAAGTATGCAAGTATG 546  
 Db 481 TAATGATTTTCACTTTGAAGTGTCACTTTGTAACCTGCAAGTATGCAAGTATG 540  
 Qy 547 AACCTGCTGGGCTATCTGCAAAAGATAATACCAACCAACCAACCAACCAACCAAC 606  
 Db 541 AACCTGCTGGGCTATCTGCAAAAGATAATCCAAACCAACCAACCAACCAACCAAC 600  
 Qy 607 CAAGCCTACAAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCAACCAAC 666  
 Db 601 CAAGCCTACAAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCAACCAAC 660  
 Qy 667 TAAACCAAGGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 697  
 Db 661 TAAATCAAGGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 691

RESULT 15  
 AAV17552  
 ID AAV17552 standard; cDNA; 15225 BP.  
 AC AAV17552;  
 XX  
 DT 20-JUL-1998 (first entry)  
 XX  
 DE Respiratory syncytial virus genome.  
 XX  
 XX RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.  
 OS  
 OS Human respiratory syncytial virus B-1.  
 PN WO9802530-A1.  
 XX  
 PD 22-JAN-1998.  
 XX  
 XX 15-JUL-1997; 97WO-US12269.  
 PF  
 XX 23-MAY-1997; 97US-0047634.  
 PR 15-JUL-1996; 96US-0021773.  
 PR 09-MAY-1997; 97US-0046141.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN;  
 PI Whitehead SS;  
 XX  
 XX WPI; 1998-110579/10.  
 DR  
 XX Attenuated respiratory syncytial virus vaccines - useful to protect  
 PT individuals against RSV infection



XX

Example 4; Page 195-202; 238pp; English.

PS

CC This is the complete nucleotide sequence of the wild-type B-1  
 CC respiratory syncytial virus (RSV). The genome is negative-sense;  
 CC the 5'-3' positive-sense sequence of D46 is provided in AAV17553.

CC

CC A novel infectious recombinant RSV comprises a RSV genome or  
 CC antigenome, a major nucleocapsid (N) protein, a nucleocapsid  
 CC phosphoprotein (P), a large polymerase protein (L), and a RNA  
 CC polymerase elongation factor, where the recombinant RSV has at  
 CC least two attenuating mutations, one of the mutations specifying a  
 CC temperature-sensitive (ts) substitution at amino acid Phe521,  
 CC Gln31, Met1169 or Tyr1321 in the RSV polymerase gene or a ts  
 CC nucleotide substitution in the gene-start sequence of gene M2.

CC

CC Also claimed are: (i) an isolated infectious RSV particle which  
 CC comprises a recombinant RSV (antigenome, N, P, and L proteins, a  
 CC RNA polymerase elongation factor, where the (anti)genome is modified;  
 CC (ii) to ablate or modulate expression of a SH, NS1, NS2 or G gene or  
 CC a cis-acting regulatory sequence; and (iii) by a termination codon  
 CC introduced within a selected gene, or by a change in sequence,  
 CC position or presence of a GS or GE transcription signal relative to  
 CC the selected gene; (2) an expression vector; and (3) an RSV strain  
 CC selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR  
 CC 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452),  
 CC cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or  
 CC B-1 cp52/2B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated  
 CC attenuated recombinant RSV and RSV particles are used in a vaccine  
 CC to stimulate the immune system of an individual to induce  
 CC protection against RSV. The expression vector of (2) is used for  
 CC the production of infectious attenuated RSV particles.

CC

XX Sequence 15225 BP; 5904 A; 2710 C; 2399 G; 4212 T; 0 other;

SQ

Query Match

Best Local Similarity 48.0%; Score 441.4; DB 19; Length 15225;

Matches 629; Conservative 0; Mismatches 291; Indels 3; Gaps 1;

QY

1 TGCAAAACATGTCCTCAAAACAGGACCAACGACCGCTAAGACACTAGAAAGACCTGGGA 60  
 4683 TGCAACCATGTCCTCAAAACAGGACCAACGACCGCTAAGACACTAGAAAGACCTGGGA 4742

DB

61 CACTCTCAATCATTTATTTATATCATATCGGGCTTATATAGTTAAATCTTAAATCTGT 120  
 4743 TACTCTCAATCATTTATTTATATCATATCGGGCTTATATAGTTAAATCTTAAATCTGT 4802

DB

121 AGCAAAATCATATTCATTTCTGGCAATGATTAATCTCAACTTCACTTAAATACAGC 180  
 4803 AGCAAAATCATATTCATTTCTGGCAATGATTAATCTCAACTTCTCTATAATTGCAGC 4862

DB

181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTACACTGCAATCATACAGA 240  
 4863 CATATATTCATATCTCTGCAATCAAAAGTTACACTAACACGGTTCAGTTCAAC 4922

QY

241 TGCAACAGCCAGATCAAGAACACACACCCCACTACTCTCAGGATCTCAGCTTGG 300  
 4923 AATTAACACCACTGTAACCAACCACTCAGGATCTCAGGATCTCAGCTTGG 4982

DB

301 AATCAGCTTCTCCATCTGTCTGAAATTAATCAATCAACCAACCACTTACTAGCTTCAAC 360  
 4983 GGTTAGCTCATCCAAACCAACCTACCAATCAACCAATCAACCAATTCAGGCCAAC 5042

QY

361 AACACGAGGATCAGTCAAGTCACACCTGCAACCCCAACCACTCAAGCTAACAACCAAC 420  
 5043 ATCAACCAACCAAGTCAGAACCAACCACTCAAGCTAACAACCACTCAAGCTAACAACCAAC 5102

DB

421 AACCCAAACACACCCAGACGCCACTTACAAACCAACCGCAACCAACCAACCAACCAAC 480  
 5103 CACTCAACACAGCAACCAACCGAGCAACCAACCGCTTAAATTCACCAACCAAC 5162

QY

481 ACCCAATATGATTTTCACTTCGAAGTGTGTTAACTTTGTACCTGTCAGCATATGACGAA 540  
 5163 ACCAAAGATGATTACCAATTTTGAAGTGTGTTAACTTTGTACCTGTCAGCATATGACGAA 5222

DB

QY 541 CAATCCAACTGCTGGGCTATCTGTCAAAAGAAATACCAACCAAAACACCAAGGAAGAAAC 600

DB 5223 CAATCACTTTTGCAAAATCTCATCTGTAAACAATACCAAGCAACAAACCAAGGAAGAAAC 5282

QY 601 CACCACCAAGCTTACAAAACCAACCAACCTTTCAA---GACAACCAACCAAAAGATCTCAAAC 657

DB 5283 AACCATCAACCCACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5342

QY 658 TCAACCACTTAACCAACCAAGGAAGTACCCACCAAGCCCAAGCCCAAGAGAGCAACCAATCAA 717

DB 5343 ACCAGCCAAACGACGAAAAAGAAACTACCAACCAACCAACCAACCAACCAACCAACCAAC 5402

QY 718 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 777

DB 5403 GACCACGAAGAGACACCAAGCACTTCAACCACTGCTCGACCAACCAACCAACCAACCAAC 5462

QY 778 ACTCAACAGTCAAAATGGAACCTTTCCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTC 837

DB 5463 ACACACAATCCACAGCAATCCCTCCACTCAACCAACCAACCAACCAACCAACCAACCAAC 5522

QY 838 TCAAGTCTCAACCAACCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 897

DB 5523 ACAACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5582

QY 898 CCAGTAGTTATTAAAAA 920

DB 5583 TGCTTAGTTATTCAAAACTACA 5605

Search completed: October 29, 2003, 22:09:33

Job time : 316.982 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 16:51:32 ; Search time 3762.15 Seconds  
(without alignments)  
10004.074 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgcgaacatgcccacaaaca.....gtagtattataaaaaaaaaa 920

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.roi.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rod.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Query |       |        | DB  | ID        | Description |
|------------|-------|-------|--------|-----|-----------|-------------|
|            | Score | Match | Length |     |           |             |
| c          | 1     | 920   | 100.0  | 920 | 6         | AI6257      |
|            |       | 920   | 100.0  | 920 | 6         | AI6258      |
|            | 3     | 920   | 100.0  | 920 | 6         | AR080406    |
|            | 4     | 920   | 100.0  | 920 | 6         | AR092530    |
|            | 5     | 920   | 100.0  | 920 | 6         | AR122885    |
|            | 6     | 920   | 100.0  | 920 | 6         | AR123540    |
| 7          | 920   | 100.0 | 920    | 6   | AR148357  |             |
| 8          | 920   | 100.0 | 920    | 6   | BD081933  |             |
| 9          | 911.8 | 99.1  | 923    | 14  | HRSNRAG   |             |
| 10         | 909.2 | 98.8  | 922    | 14  | HRSVGL16  |             |
| 11         | 901.6 | 98.0  | 917    | 14  | RSHGLYG   |             |
| 12         | 894   | 97.2  | 894    | 6   | AR080424  |             |
| 13         | 894   | 97.2  | 894    | 6   | AR092548  |             |
| 14         | 853.4 | 92.8  | 8510   | 14  | RSHICE    |             |
| 15         | 853.4 | 92.8  | 15222  | 6   | AR093219  |             |
| 16         | 853.4 | 92.8  | 15222  | 6   | BD091237  |             |
| 17         | 853.4 | 92.8  | 15222  | 14  | HRU50362  |             |
| 18         | 853.4 | 92.8  | 15222  | 14  | HRU50363  |             |
| 19         | 853.4 | 92.8  | 15222  | 14  | HRU63644  |             |
| 20         | 853.4 | 92.8  | 15222  | 14  | RSHSEQ    |             |
| 21         | 853.4 | 92.8  | 15223  | 6   | AR089137  |             |
| 22         | 853.4 | 92.8  | 15223  | 14  | AF035006  |             |
| 23         | 848.2 | 92.2  | 914    | 14  | PARSENVG  |             |
| 24         | 848.2 | 92.2  | 918    | 6   | AX047075  |             |
| 25         | 837.8 | 91.1  | 897    | 6   | AX339011  |             |
| 26         | 837.2 | 91.0  | 15190  | 14  | HRU39662  |             |
| 27         | 835.6 | 90.8  | 15191  | 14  | RSU39661  |             |
| 28         | 820.4 | 89.2  | 918    | 14  | AF065406  |             |
| 29         | 814   | 88.5  | 918    | 14  | AF065405  |             |
| 30         | 811.6 | 88.2  | 922    | 14  | HRSCVGL1  |             |
| 31         | 811.6 | 88.2  | 922    | 14  | HRSCVGA   |             |
| 32         | 810   | 88.0  | 922    | 14  | HRSCVG3   |             |
| 33         | 810   | 88.0  | 922    | 14  | HRSCVGC   |             |
| 34         | 810   | 88.0  | 922    | 14  | HRSVGL11  |             |
| 35         | 810   | 88.0  | 922    | 14  | HRSVGL12  |             |
| 36         | 808.4 | 87.9  | 922    | 14  | HRSVGGLY1 |             |
| 37         | 808.4 | 87.9  | 922    | 14  | HRSVGGLY2 |             |
| 38         | 807.2 | 87.7  | 924    | 14  | HRSVGL14  |             |
| 39         | 806.8 | 87.7  | 922    | 14  | HRSVGL2   |             |
| 40         | 806.8 | 87.7  | 922    | 14  | HRSVGL9   |             |
| 41         | 805.2 | 87.5  | 922    | 14  | AF065256  |             |
| 42         | 805.2 | 87.5  | 922    | 14  | HRSVGL10  |             |
| 43         | 805.2 | 87.5  | 922    | 14  | HRSVGL13  |             |
| 44         | 805.2 | 87.5  | 922    | 14  | HRSVGL3   |             |
| 45         | 803.6 | 87.3  | 922    | 14  | AF065255  |             |

# ALIGNMENTS

RESULT 1  
AI6257  
LOCUS AI6257 Synthetic RSV G gene (seq ID No: 7).  
DEFINITION AI6257  
ACCESSION AI6257  
VERSION AI6257.1 GI:640933  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 920)  
AUTHORS  
TITLE CHIMERIC IMMUNOGENS  
JOURNAL Patent: WO 9314207-A 7 22-JUL-1993;  
FEATURES Location/Qualifiers  
linear PAT 03-OCT-1994  
920 bp DNA

```

source
1. .920
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
8. .904
/gene="RSV G"
8. .904
/gene="RSV G"
8. .904
/codon_start=1
/transl_table=11
/protein_id="CAA01273.1"
/db_xref="GI:640934"
/translation="MSKNDQORTAKTLEKWTDLHLLFTSSGLYKLNKLSVAQITLS
ILAMTISLLIATIFIASANHKVLTITAIQDATSOIKNTFTYLTQDPQLGISFS
NLSEITSTQITLSTFTGKSNLQPTTKNTTFTQTPSKPTKQKPNKPN
NDHPFVFNFPVPCISCSNNPTKMAICKRIPNKPKGKTKTKPKTKTKKDLKQ
TKPKKEVPTKTEPTINTTKNTITLITLLTNNTTGNPKLTSQMEIFHSTSSBGNLSP
SQVSTTSBHPSPSPNTTQ"
BASE COUNT      380 a   290 c   95 g   155 t
ORIGIN
Query Match      100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3e-192;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 TGCACACATGTCACAAAACAGGACCAACGACCGCTTAAGACACATAGAAAAGACCTGGGA 60
DB      1 TGCACACATGTCACAAAACAGGACCAACGACCGCTTAAGACACATAGAAAAGACCTGGGA 60
QY      61 CACTCTCAATCATTTATTATTATCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120
DB      61 CACTCTCAATCATTTATTATTATCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120
QY      121 AGCACAAATCACATTAATCTCTGGCAATGATATCTCAACTTCACCTTAAATTAATACAGC 180
DB      121 AGCACAAATCACATTAATCTCTGGCAATGATATCTCAACTTCACCTTAAATTAATACAGC 180
QY      181 CATCATATTCATAGCTCTGGCAAAACCAACAAAGTCACACTAAACACTGCAATCATACAAGA 240
DB      181 CATCATATTCATAGCTCTGGCAAAACCAACAAAGTCACACTAAACACTGCAATCATACAAGA 240
QY      241 TGCACACAGCCAGATCAGACACACACCCACACATCTCTCAGTCTCAGGATCTCAGCTTGG 300
DB      241 TGCACACAGCCAGATCAGACACACACCCACACATCTCTCAGTCTCAGGATCTCAGCTTGG 300
QY      301 AATCAGCTTCTCAATCTGTCTGAAATTAATCATCAACACCAACCACTACTAGCTTCAAC 360
DB      301 AATCAGCTTCTCAATCTGTCTGAAATTAATCATCAACACCAACCACTACTAGCTTCAAC 360
QY      361 AACACGAGAGTCAAGTCAACCTGCAACCCACACACAGTCAAGCTAAGAAACCAACCAAC 420
DB      361 AACACGAGAGTCAAGTCAACCTGCAACCCACACACAGTCAAGCTAAGAAACCAACCAAC 420
QY      421 AACCACAAACACACACCCAGCAGCCACCTACAAACCAACGACCAACCAACCAACCAAC 480
DB      421 AACCACAAACACACACCCAGCAGCCACCTACAAACCAACGACCAACCAACCAACCAAC 480
QY      481 ACCCAATATGATTTTTCATCTCGAAGTGTTTAACTTTGTACCTGCGAGCATATGACGAA 540
DB      481 ACCCAATATGATTTTTCATCTCGAAGTGTTTAACTTTGTACCTGCGAGCATATGACGAA 540
QY      541 CAATCCAGCTCTGCTGGGTATCTGCAAAAGAAATACCAACCAACCAACCAACCAACCAAC 600
DB      541 CAATCCAGCTCTGCTGGGTATCTGCAAAAGAAATACCAACCAACCAACCAACCAACCAAC 600
QY      601 CACCACCAAGCTTACAAAACCAACCACTTCAAGACACCAACCAACCAACCAACCAACCTCA 660
DB      601 CACCACCAAGCTTACAAAACCAACCACTTCAAGACACCAACCAACCAACCAACCAACCTCA 660
QY      661 AACCACTAAACCAAGGAGTACCAACCAACCAAGCCCAACGAGAGGACCAACCACTCAAC 720
DB      661 AACCACTAAACCAAGGAGTACCAACCAACCAAGCCCAACGAGAGGACCAACCACTCAAC 720

```

```
Db 500 AACCCAAACACAAACCCAGAGCCCACTACAAAACAAAGCCCAAAACAAACCAACCAAA 441
Qy 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGTACCTCGAGCATATGAGCAA 540
Db 440 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGTACCTCGAGCATATGAGCAA 381
Qy 541 CAATCAACCTGCTGGGCTATCTGCAAAAGATATCAAAACAAACCAAGGAAAGAAAC 600
Db 380 CAATCAACCTGCTGGGCTATCTGCAAAAGATATCAAAACAAACCAAGGAAAGAAAC 321
Qy 601 CACCAAGAGCTTACAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCAAC 660
Db 320 CACCAAGAGCTTACAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCAAC 261
Qy 661 AACCACTAAACCAAGAGAGTACCCACCAAGCCACAGAGAGCCCAACCATCAAC 720
Db 260 AACCACTAAACCAAGAGAGTACCCACCAAGCCACAGAGAGCCCAACCATCAAC 201
Qy 721 CACCAAAACAAACATCAACACTTACTGCTCAACCAACCAACCAACCAACCAAC 780
Db 200 CACCAAAACAAACATCAACACTTACTGCTCAACCAACCAACCAACCAACCAAC 141
Qy 781 CACAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 840
Db 140 CACAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 81
Qy 841 AGTCTCCAAACATCCGAGAGCCCAATCAACCTTCAACCTTCCGAGGCAATCTTAAG 900
Db 80 AGTCTCCAAACATCCGAGAGCCCAATCAACCTTCAACCTTCCGAGGCAATCTTAAG 21
Qy 901 GTAGTTATTAATAAAAAA 920
Db 20 GTAGTTATTAATAAAAAA 1

RESULT 3
AR080406
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 920)
AUTHORS
TITLE
Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus
JOURNAL
Patent: US 5968776-A 7 19-OCT-1999;
FEATURES
Location/Qualifiers
1..920
source
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3e-192;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCACCAATGTCACAAACCAAGGAGCCACCGCTAAGACACTAGAAAGACCTGGGA 60
Db 1 TGCACCAATGTCACAAACCAAGGAGCCACCGCTAAGACACTAGAAAGACCTGGGA 60
Qy 61 CACTCTCAATCATTATTTATCATATCATCGGGCTTATATAAGTTAAATCTTAATCTGT 120
Db 61 CACTCTCAATCATTATTTATCATATCATCGGGCTTATATAAGTTAAATCTTAATCTGT 120
Qy 121 AGCACAATCATTATTCATTTCTGGCAATGATAATCTCAACTTCACTTATAATTACAGC 180
Db 121 AGCACAATCATTATTCATTTCTGGCAATGATAATCTCAACTTCACTTATAATTACAGC 180
```

```
Qy 181 CATCATATTTAGCTTCGGCCAAACCAAGTCACTAACTGCAATCATACAAGA 240
Db 181 CATCATATTTAGCTTCGGCCAAACCAAGTCACTAACTGCAATCATACAAGA 240
Qy 241 TGCACCAAGCCAGATCAAGAACACAAACCCCAATACCTCACTCAGGATCTCTCAGCTTGG 300
Db 241 TGCACCAAGCCAGATCAAGAACACAAACCCCAATACCTCACTCAGGATCTCTCAGCTTGG 300
Qy 301 AATCAGCTTCTCCAAATCTGTCTGAAATTTATCATCAAAACCAACCACTACTAGCTTCAAC 360
Db 301 AATCAGCTTCTCCAAATCTGTCTGAAATTTATCATCAAAACCAACCACTACTAGCTTCAAC 360
Qy 361 AACCAACGAGTCAAGTCAACCTGCAACCAACCAACCAACCAACCAACCAACCAAC 420
Db 361 AACCAACGAGTCAAGTCAACCTGCAACCAACCAACCAACCAACCAACCAACCAAC 420
Qy 421 AACCAACCAACCAACGAGGAGCCACTACAAAACCAACCAACCAACCAACCAACCAAC 480
Db 421 AACCAACCAACCAACGAGGAGCCACTACAAAACCAACCAACCAACCAACCAACCAAC 480
Qy 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGTACCTCGAGCATATGAGCAA 540
Db 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGTACCTCGAGCATATGAGCAA 540
Qy 541 CAATCCCAACCTGCTGGGCTATCTGCAAAAGATACCAAAACCAAGGAAAGAAAC 600
Db 541 CAATCCCAACCTGCTGGGCTATCTGCAAAAGATACCAAAACCAAGGAAAGAAAC 600
Qy 601 CACCAACCAAGCTTCAAAAAACCAACCTTCAAGACCAACCAAAACCAAGGAAAGAAAC 660
Db 601 CACCAACCAAGCTTCAAAAAACCAACCTTCAAGACCAACCAAAACCAAGGAAAGAAAC 660
Qy 661 AACCACTAAACCAAGGAGTACCCACCAACCAACCAACCAACCAACCAACCAAC 720
Db 661 AACCACTAAACCAAGGAGTACCCACCAACCAACCAACCAACCAACCAACCAAC 720
Qy 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780
Db 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780
Qy 781 CACAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 840
Db 781 CACAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 840
Qy 841 AGTCTCCAAACATCCGAGAGCCCAACCTTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 900
Db 841 AGTCTCCAAACATCCGAGAGCCCAACCTTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 900
Qy 901 GTAGTTATTAATAAAAAA 920
Db 901 GTAGTTATTAATAAAAAA 920

RESULT 4
AR092530
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 920)
AUTHORS
TITLE
Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus
JOURNAL
Patent: US 5998169-A 7 07-DEC-1999;
FEATURES
Location/Qualifiers
1..920
source
BASE COUNT 380 a 290 c 95 g 155 t
```

| ORIGIN  |                       |   |           |         |    |
|---------|-----------------------|---|-----------|---------|----|
|         | Query Match           | 100.0%; Score 920; DB 6; Length 920;                            |           |         |    |
|         | Best Local Similarity | 100.0%; Pred. No. 3e-192;                                       |           |         |    |
| Matches | 920; Conservative     | 0; Mismatches   | 0; Indels | 0; Gaps | 0; |
| Qy      | 1                     | TGCAAAATGTCCAAAAACAAGGACCAAAGCACCAGCTTAAGACACTAGAAAAGACTGGGA    | 60        |         |    |
| Dd      | 1                     | TGCAAAATGTCCAAAAACAAGGACCAAAGCACCAGCTTAAGACACTAGAAAAGACTGGGA    | 60        |         |    |
| Qy      | 61                    | CACCTCTCAATCATTTATTTCATATCATCGGGCTTATTAAGTTAATCTTAAATCTGT       | 120       |         |    |
| Dd      | 61                    | CACCTCTCAATCATTTATTATTCATATCATCGGGCTTATTAAGTTAATCTTAAATCTGT     | 120       |         |    |
| Qy      | 121                   | AGCAAAATCACATTTATCCATTTCTGGCAATGATAATCTCAACTTCACCTTATAATTTACAGC | 180       |         |    |
| Dd      | 121                   | AGCAAAATCACATTTATCCATTTCTGGCAATGATAATCTCAACTTCACCTTATAATTTACAGC | 180       |         |    |
| Qy      | 181                   | CATCATATTATAGCTTGCGAAACCAACAAAGTGCACATAACAATGCAATCAFAACAAGA     | 240       |         |    |
| Dd      | 181                   | CATCATATTATAGCTTGCGAAACCAACAAAGTGCACATAACAATGCAATCAFAACAAGA     | 240       |         |    |
| Qy      | 241                   | TGCACAGGCCAGATCAGAACACAAACCCACATACCTCTCACTCAGGATCTCTCAGCTTGG    | 300       |         |    |
| Dd      | 241                   | TGCACAGGCCAGATCAAGAACACAAACCCACATACCTCTCACTCAGGATCTCTCAGCTTGG   | 300       |         |    |
| Qy      | 301                   | AATCAGCTTCTCCAATCTGTCTGAAATPACATCAACAAACCAACCACTACTAGCTTCAAC    | 360       |         |    |
| Dd      | 301                   | AATCAGCTTCTCCAATCTGTCTGAAATPACATCAACAAACCAACCACTACTAGCTTCAAC    | 360       |         |    |
| Qy      | 361                   | AACACCGAGGTCAAGTCAAACTCTGCAACCCACAAGTCAAGACTTAAAAACAACAACAC     | 420       |         |    |
| Dd      | 361                   | AACACCGAGGTCAAGTCAAACTCTGCAACCCACAAGTCAAGACTTAAAAACAACAACAC     | 420       |         |    |
| Qy      | 421                   | AAACCAACACAAACCCAGACGCCACTTACAAAACACGCCAAACCAACCAACCAACAA       | 480       |         |    |
| Dd      | 421                   | AAACCAACACAAACCCAGACGCCACTTACAAAACACGCCAAACCAACCAACCAACAA       | 480       |         |    |
| Qy      | 481                   | ACCCAATGAATTTTCACTTTCGAAGTGTAACTTTGTACCTCTGCAGCATATGCAGCA       | 540       |         |    |
| Dd      | 481                   | ACCCAATGAATTTTCACTTTCGAAGTGTAACTTTGTACCTCTGCAGCATATGCAGCA       | 540       |         |    |
| Qy      | 541                   | CAATCCAACTGTGGGCTATCTGCAAAAGAATACCAAAACAAAAACCAAGAAAGAAAC       | 600       |         |    |
| Dd      | 541                   | CAATCCAACTGTGGGCTATCTGCAAAAGAATACCAAAACAAAAACCAAGAAAGAAAC       | 600       |         |    |
| Qy      | 601                   | CACACCAAGGCTACAAAAAACCAACCTTCAGACCAACCAAAAAAGATCTCAAACTCA       | 660       |         |    |
| Dd      | 601                   | CACACCAAGGCTACAAAAAACCAACCTTCAGACCAACCAAAAAAGATCTCAAACTCA       | 660       |         |    |
| Qy      | 661                   | AACCACTAAACCAAAAGGAAGTACCAACCAACCAAGCCCAAGAGAGCCAAACATCAAC      | 720       |         |    |
| Dd      | 661                   | AACCACTAAACCAAAAGGAAGTACCAACCAACCAAGCCCAAGAGAGCCAAACATCAAC      | 720       |         |    |
| Qy      | 721                   | CACCAAAACAAACATCAACAATTACATGCTCTCAACCAACCAACCAAGAAATCCAAAAC     | 780       |         |    |
| Dd      | 721                   | CACCAAAACAAACATCAACAATTACATGCTCTCAACCAACCAACCAAGAAATCCAAAAC     | 780       |         |    |
| Qy      | 781                   | CACAAGTCAATGGAACCTTCCACTCACTCTCCGAAGGCAATCTAAGCCCTTCTCA         | 840       |         |    |
| Dd      | 781                   | CACAAGTCAATGGAACCTTCCACTCACTCTCCGAAGGCAATCTAAGCCCTTCTCA         | 840       |         |    |
| Qy      | 841                   | AGTCTCCCAACATCCGAGCACCATCAACACCTCTCTCCACCCCAACACAAACGCA         | 900       |         |    |
| Dd      | 841                   | AGTCTCCCAACATCCGAGCACCATCAACACCTCTCTCCACCCCAACACAAACGCA         | 900       |         |    |
| Qy      | 901                   | GTAGTTATTAATAAAAAAAAA 920                                       |           |         |    |
| Dd      | 901                   | GTAGTTATTAATAAAAAAAAA 920                                       |           |         |    |

RESULT 5  
AR122885

|            |   |                 |  |           |                 |
|------------|---|-----------------|--|-----------|-----------------|
| LOCUS      | AR122885                                | 920 bp          | DNA  | linear    | PAT 16-MAY-2001 |
| DEFINITION | Sequence 7 from patent US 6168786.      |                 |  |           |                 |
| ACCESSION  | AR122885                                |                 |  |           |                 |
| VERSION    | AR122885.1                              | GI:14107851     |  |           |                 |
| KEYWORDS   | Unknown.                                |                 |  |           |                 |
| SOURCE     | Unknown.                                |                 |  |           |                 |
| ORGANISM   | Unclassified.                           |                 |  |           |                 |
| REFERENCE  | 1 (bases 1 to 920)                      |                 |  |           |                 |
| AUTHORS    | Klein,M.H., Du,R.-p. and Ewaszshyn,M.E. |                 |  |           |                 |
| TITLE      | Chimeric immunogens                     |                 |  |           |                 |
| JOURNAL    | Patent: US 6168786-A 7 02-JAN-2001;     |                 |  |           |                 |
| FEATURES   | Location/Qualifiers                     |                 |  |           |                 |
| source     | 1..920                                  |                 |  |           |                 |
| BASE COUNT | 380 a                                   | 290 c           | 95 g   | 155 t     |                 |
| ORIGIN     |   |                 |  |           |                 |
|            | Query Match                             | 100.0%;         | Score 920;   | DB 6;     | Length 920;     |
|            | Best Local Similarity                   | 100.0%;         | Pred. No. 3e-192;  |           |                 |
|            | Matches 920;                            | Conservative 0; | Mismatches 0;  | Indels 0; | Gaps 0;         |
| Qy         | 1                                       | TGCAAA          | CATGTCCAAAACAGGACCAAGCGACCGCTTAAGACACTAGAAAGACCTGGGA         | 60        |                 |
| Db         | 1                                       | TGCAAA          | CATGTCCAAAACAGGACCAAGCGACCGCTTAAGACACTAGAAAGACCTGGGA         | 60        |                 |
| Qy         | 61                                      | CACCTC          | CAATCATTTATTATTATATCATATCATATCGGGCTTATATAAGTTAAATCTTAAATCTGT | 120       |                 |
| Db         | 61                                      | CACCTC          | CAATCATTTATTATTATATCATATCATATCGGGCTTATATAAGTTAAATCTTAAATCTGT | 120       |                 |
| Qy         | 121                                     | AGCAAA          | AATCAATTTCCATTTCTGGCAATGATTAATCTCAACTTCACTTATAATTACAGC       | 180       |                 |
| Db         | 121                                     | AGCAAA          | AATCAATTTCCATTTCTGGCAATGATTAATCTCAACTTCACTTATAATTACAGC       | 180       |                 |
| Qy         | 181                                     | CATCAT          | TATCATAGCTTCGGCAACCAACCAAGTCACTAACTGCATCATACAAGA             | 240       |                 |
| Db         | 181                                     | CATCAT          | TATCATAGCTTCGGCAACCAACCAAGTCACTAACTGCATCATACAAGA             | 240       |                 |
| Qy         | 241                                     | TGCAAC          | AGCCAGATCAAGAACACCAACCCCAACTACTCTCAGGATCTCTCAGCTTGG          | 300       |                 |
| Db         | 241                                     | TGCAAC          | AGCCAGATCAAGAACACCAACCCCAACTACTCTCAGGATCTCTCAGCTTGG          | 300       |                 |
| Qy         | 301                                     | AATCAG          | CTTCTCCAATCTGTCTGAATTTACATCAAAACCAACCACTACTAGCTTCAAC         | 360       |                 |
| Db         | 301                                     | AATCAG          | CTTCTCCAATCTGTCTGAATTTACATCAAAACCAACCACTACTAGCTTCAAC         | 360       |                 |
| Qy         | 361                                     | AACACG          | AGGATCAAGTCAAACTTCGAACCCCAACAGTCAAGCTAAAAACACACAAC           | 420       |                 |
| Db         | 361                                     | AACACG          | AGGATCAAGTCAAACTTCGAACCCCAACAGTCAAGCTAAAAACACACAAC           | 420       |                 |
| Qy         | 421                                     | AACCAAA         | ACACACGACGACGACCACTACAAAACACGCGCAAAACCAACCAACAA              | 480       |                 |
| Db         | 421                                     | AACCAAA         | ACACACGACGACGACCACTACAAAACACGCGCAAAACCAACCAACAA              | 480       |                 |
| Qy         | 481                                     | ACCCAAT         | AATGATTTTCACTTCGAAGTGTTTAACTTTTGACCCCTGCAGCATATGCAGCAA       | 540       |                 |
| Db         | 481                                     | ACCCAAT         | AATGATTTTCACTTCGAAGTGTTTAACTTTTGACCCCTGCAGCATATGCAGCAA       | 540       |                 |
| Qy         | 541                                     | CAATCC          | AACTCTGGGGTATCTGCAAAAGATACAAAACCAAAAACACGAGGAAGAAAC          | 600       |                 |
| Db         | 541                                     | CAATCC          | AACTCTGGGGTATCTGCAAAAGATACAAAACCAAAAACACGAGGAAGAAAC          | 600       |                 |
| Qy         | 601                                     | CACCAC          | CAGGCTACAAAAAACCAACTTCACAGACACCAAAAAGATCTCAAACTCA            | 660       |                 |
| Db         | 601                                     | CACCAC          | CAGGCTACAAAAAACCAACTTCACAGACACCAAAAAGATCTCAAACTCA            | 660       |                 |
| Qy         | 661                                     | AACCACT         | AAACCAAGGAAGTACCCACCAACGAGCCCAAGAGAGCAACCATCAACAC            | 720       |                 |
| Db         | 661                                     | AACCACT         | AAACCAAGGAAGTACCCACCAACGAGCCCAAGAGAGCAACCATCAACAC            | 720       |                 |
| Qy         | 721                                     | CACCAAA         | CAAAACATCTCAACTACATGTGTCACCAACACCAACCAAGGAATCCAAACT          | 780       |                 |
| Db         | 721                                     | CACCAAA         | CAAAACATCTCAACTACATGTGTCACCAACACCAACCAAGGAATCCAAACT          | 780       |                 |

781 CACAAGTCAATGGAACCTTTCACTCACTCACTCTCCGAGGGAATCAAGCCCTTCTCA 840  
Db CACAAGTCAATGGAACCTTTCACTCACTCTCCGAGGGAATCAAGCCCTTCTCA 840  
781 CACAAGTCAATGGAACCTTTCACTCACTCTCCGAGGGAATCAAGCCCTTCTCA 840  
841 AGTCTCCACAAATCCGAGCACCACATCAACCTCTATCTCCACCCCAACACAGCCCA 900  
Db AGTCTCCACAAATCCGAGCACCACATCAACCTCTATCTCCACCCCAACACAGCCCA 900  
901 GTAGTTATTAAAAA 920  
Db GTAGTTATTAAAAA 920

RESULT 6  
AR123540  
LOCUS AR123540 920 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 7 from patent US 6171783.  
ACCESSION AR123540  
VERSION AR123540.1 GI:14108901  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewaszyshyn, M.E.  
TITLE Infection detection method using chimeric protein  
JOURNAL Patent: US 6171783-A 7 09-JAN-2001;  
FEATURES Location/Qualifiers  
source  
1. .920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN

Query Match 100.0%; Score 920; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3e-192;  
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGCACAACTGTCCAAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
Db 1 TGCACAACTGTCCAAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
61 CACTCTCAATCAATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTCT 120  
Db 61 CACTCTCAATCAATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTCT 120  
121 AGCAGAAATCAATATTCATATTCATCGGCAATGATAATCTCACTTCACTTATTAATACAG 180  
Db 121 AGCAGAAATCAATATTCATATTCATCGGCAATGATAATCTCACTTCACTTATTAATACAG 180  
181 CATCATATTCATAGCTTCGGCAACCAACAAAGTCACTAACTGCAATCATACAAGA 240  
Db 181 CATCATATTCATAGCTTCGGCAACCAACAAAGTCACTAACTGCAATCATACAAGA 240  
241 TGCACAAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300  
Db 241 TGCACAAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300  
301 AATCAGCTTCTCAATCTGTCTGAAATTAATCAATCAAAACCAACATTAATAGCTTCAAC 360  
Db 301 AATCAGCTTCTCAATCTGTCTGAAATTAATCAATCAAAACCAACATTAATAGCTTCAAC 360  
361 AACACAGAGGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420  
Db 361 AACACAGAGGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420  
421 AACCCAAACACAAACCCAGAGCCCACTACAAAACAAAGCCCAAAACCAACCAACAA 480  
Db 421 AACCCAAACACAAACCCAGAGCCCACTACAAAACAAAGCCCAAAACCAACCAACAA 480  
481 ACCCAATATGATTTTCACTTCAAGTGTATTTTGTACCTTGTACCTGAGCATATGAGCAA 540  
Db 481 ACCCAATATGATTTTCACTTCAAGTGTATTTTGTACCTTGTACCTGAGCATATGAGCAA 540

541 CAATCCAACTGTCTGGGCTATCTGCAAAAGATACCAACAAAAACCGAGAAAGAAAC 600  
Db 541 CAATCCAACTGTCTGGGCTATCTGCAAAAGATACCAACAAAAACCGAGAAAGAAAC 600  
601 CACCAACAAAGCCCTTACAAAAAACCAACCTTCAAGACCAACCAAAAAAGATCTCAAACTCA 660  
Db 601 CACCAACAAAGCCCTTACAAAAAACCAACCTTCAAGACCAACCAAAAAAGATCTCAAACTCA 660  
661 AACCACTAAACCAAAAGAGTACCCACCAACGAGCCCAACAGAGCCCAACCACTCAACAC 720  
Db 661 AACCACTAAACCAAAAGAGTACCCACCAACGAGCCCAACAGAGCCCAACCACTCAACAC 720  
721 CACCAAAACAAACATCACTACACTGCTGCTACCAACCAACCAACAGGAAATCCAAAACCT 780  
Db 721 CACCAAAACAAACATCACTACACTGCTGCTACCAACCAACCAACAGGAAATCCAAAACCT 780  
781 CACAAGTCAATGGAACCTTCCACTCAACCTTCCCGAGGCAATCTTAAGCCCTTCTCA 840  
Db 781 CACAAGTCAATGGAACCTTCCACTCAACCTTCCCGAGGCAATCTTAAGCCCTTCTCA 840  
841 AGTCTCCACAAATCCGAGCACCACATCAACCTCTCATCTCCACCAACCAACAGCCCA 900  
Db 841 AGTCTCCACAAATCCGAGCACCACATCAACCTCTCATCTCCACCAACCAACAGCCCA 900  
901 GTAGTTATTAAAAA 920  
Db 901 GTAGTTATTAAAAA 920

RESULT 7  
AR148357  
LOCUS AR148357 920 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 7 from patent US 6225091.  
ACCESSION AR148357  
VERSION AR148357.1 GI:15112447  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewaszyshyn, M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
JOURNAL Patent: US 6225091-A 7 01-MAY-2001;  
FEATURES Location/Qualifiers  
source  
1. .920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN

Query Match 100.0%; Score 920; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3e-192;  
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGCACAACTGTCCAAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
Db 1 TGCACAACTGTCCAAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
61 CACTCTCAATCAATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120  
Db 61 CACTCTCAATCAATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120  
121 AGCAGAAATCAATATTCATATTCATCGGCAATGATAATCTCACTTCACTTATTAATACAG 180  
Db 121 AGCAGAAATCAATATTCATATTCATCGGCAATGATAATCTCACTTCACTTATTAATACAG 180  
181 CATCATATTCATAGCTTCGGCAACCAACAAAGTCACTAACTGCAATCATACAAGA 240  
Db 181 CATCATATTCATAGCTTCGGCAACCAACAAAGTCACTAACTGCAATCATACAAGA 240  
241 TGCACAAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300  
Db 241 TGCACAAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300  
301 AATCAGCTTCTCAATCTGTCTGAAATTAATCAATCAAAACCAACATTAATAGCTTCAAC 360  
Db 301 AATCAGCTTCTCAATCTGTCTGAAATTAATCAATCAAAACCAACATTAATAGCTTCAAC 360  
361 AACACAGAGGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420  
Db 361 AACACAGAGGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420  
421 AACCCAAACACAAACCCAGAGCCCACTACAAAACAAAGCCCAAAACCAACCAACAA 480  
Db 421 AACCCAAACACAAACCCAGAGCCCACTACAAAACAAAGCCCAAAACCAACCAACAA 480  
481 ACCCAATATGATTTTCACTTCAAGTGTATTTTGTACCTTGTACCTGAGCATATGAGCAA 540  
Db 481 ACCCAATATGATTTTCACTTCAAGTGTATTTTGTACCTTGTACCTGAGCATATGAGCAA 540

Db 241 TGCACAGCCAGATCAAGACACCAACCCCAACATACCTCCTCAGGATCTCAGCTTGG 300  
QY 301 AATCAGCTTCTCAATCTGTCTGAATTAATACACAAACCAACCAACATAGCTTCAAC 360  
Db 301 AATCAGCTTCTCAATCTGTCTGAATTAATACACAAACCAACCAACATAGCTTCAAC 360  
QY 361 AACACAGGAGTCAAGTCAACCTGCAACCCCAACAGTCAAGCTAATTAACCAACAC 420  
Db 361 AACACAGGAGTCAAGTCAACCTGCAACCCCAACAGTCAAGCTAATTAACCAACAC 420  
QY 421 AACCCAAACACACACCAAGCAGCCACTACAAACCAACGCGCAAAACCAACCAACAA 480  
Db 421 AACCCAAACACACACCAAGCAGCCACTACAAACCAACGCGCAAAACCAACCAACAA 480  
QY 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
Db 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
QY 541 CAATCCAACTGTGGCTATCTGCAAGAAATACCAACCAAAACCAAGGAAAGAAC 600  
Db 541 CAATCCAACTGTGGCTATCTGCAAGAAATACCAACCAAAACCAAGGAAAGAAC 600  
QY 601 CACCAACAGCTTACAAAAACCAACCTTCAAGACCAACCAAAACCAAGCTCTCAACCTCA 660  
Db 601 CACCAACAGCTTACAAAAACCAACCTTCAAGACCAACCAAAACCAAGCTCTCAACCTCA 660  
QY 661 AACCACTAAACCAAGGAGTACCCACCAACCAAGCCCAACCAAGGAGCCCAACCAAC 720  
Db 661 AACCACTAAACCAAGGAGTACCCACCAACCAAGCCCAACCAAGGAGCCCAACCAAC 720  
QY 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
Db 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840  
Db 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840  
QY 841 AGTCTCCAAACATCCGAGCACCCTACCAACCTCTCCACCCCAACCAACCAACCA 900  
Db 841 AGTCTCCAAACATCCGAGCACCCTACCAACCTCTCCACCCCAACCAACCAACCA 900  
QY 901 GTAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 920  
Db 901 GTAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 920

RESULT 8  
BD081933  
LOCUS  
DEFINITION  
Nucleic acid vaccines encoding G protein of respiratory syncytial virus.  
ACCESSION  
BD081933  
VERSION  
BD081933.1 GI:22627543  
KEYWORDS  
JP 2001512662-A/1.  
SOURCE  
Respiratory syncytial virus  
ORGANISM  
Respiratory syncytial virus  
Virus; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE  
1. (bases 1 to 920)  
Li, X., Sambhara, S. and Klein, M.H.  
Nucleic acid vaccines encoding G protein of respiratory syncytial  
TITLE  
Patent: JP 2001512662-A 1 28-AUG-2001;  
JOURNAL  
CONNAUGHT LABORATORIES LTD  
COMMENT  
OS Respiratory syncytial virus  
PN JP 2001512662-A/1  
PD 28-AUG-2001  
PF 16-JUL-1998 JP 2000503216  
PR 18-JUL-1997 US 08/896442  
PI XIAOMAO LI, SURYAPRAKESH SAMBHARA, MICHEL H KLEIN PC  
C12N15/09, A61K48/00, C07K16/10, C12P21/08, C12Q1/68, G01N33/53, PC  
C12N15/00

CC Nucleic acid vaccines encoding G protein of respiratory syncytial virus  
FH Key Location/Qualifiers  
FT CDS (7)..(901).  
FEATURES  
source  
Location/Qualifiers  
1..920  
/organism="Respiratory syncytial virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:112814"  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN  
Query Match 100.0%; Score 920; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3e-192;  
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGCACACATGTCCAAAAACAGGACCAACGCGCTAAGACACACTAGAAAAAGACCTGGCA 60  
Db 1 TGCACACATGTCCAAAAACAGGACCAACGCGCTAAGACACACTAGAAAAAGACCTGGCA 60  
QY 61 CACTCTCAATCATTTTATTTATTTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT 120  
Db 61 CACTCTCAATCATTTTATTTATTTATTTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT 120  
QY 121 AGCACAAATCATTATTCATTTCTGCAATGATATCTCAACTTCACTTCACTTATTAATTACAG 180  
Db 121 AGCACAAATCATTATTCATTTCTGCAATGATATCTCAACTTCACTTCACTTATTAATTACAG 180  
QY 181 CATCATATTTCATAGCTTCGGCAAAACCAACAAAGTCACTAAACAACTGCAATCATACAAGA 240  
Db 181 CATCATATTTCATAGCTTCGGCAAAACCAACAAAGTCACTAAACAACTGCAATCATACAAGA 240  
QY 241 TGCACACGCGAGTCAAGAACACAAACCCCAACATACCTCCTCAGGATCTCAGCTTGG 300  
Db 241 TGCACACGCGAGTCAAGAACACAAACCCCAACATACCTCCTCAGGATCTCAGCTTGG 300  
QY 301 AATCAGCTTCTCCATCTCTCTGCAATTTATCATCAACCAACCAACCACTACTAGCTTCAAC 360  
Db 301 AATCAGCTTCTCCATCTCTCTGCAATTTATCATCAACCAACCAACCACTACTAGCTTCAAC 360  
QY 361 AACACAGGAGTCAAGTCAAAACCTGCAACCCCAACCAAGTCAAGCTTAAAAACCAACAA 420  
Db 361 AACACAGGAGTCAAGTCAAAACCTGCAACCCCAACCAAGTCAAGCTTAAAAACCAACAA 420  
QY 421 AACCCAAACACAAACCCAGCAAGCCCACTACAAACAAACGCGCAAAACCAACCAACAA 480  
Db 421 AACCCAAACACAAACCCAGCAAGCCCACTACAAACAAACGCGCAAAACCAACCAACAA 480  
QY 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
Db 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
QY 541 CAATCCAACTGTGGCTATCTGCAAGAAATACCAACCAAAACCAAGGAAAGAAC 600  
Db 541 CAATCCAACTGTGGCTATCTGCAAGAAATACCAACCAAAACCAAGGAAAGAAC 600  
QY 601 CACCAACAGCTTACAAAAACCAACCTTCAAGACCAACCAAAACCAAGCTCTCAACCTCA 660  
Db 601 CACCAACAGCTTACAAAAACCAACCTTCAAGACCAACCAAAACCAAGCTCTCAACCTCA 660  
QY 661 AACCACTAAACCAAGGAGTACCCACCAACCAAGCCCAACCAAGGAGCCCAACCAAC 720  
Db 661 AACCACTAAACCAAGGAGTACCCACCAACCAAGCCCAACCAAGGAGCCCAACCAAC 720  
QY 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
Db 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840  
Db 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840  
QY 841 AGTCTCCAAACATCCGAGCACCCTACCAACCTCTCCACCCCAACCAACCAACCA 900



| Db                    | 841   | AGTCTCCACAAATCCGAGCACC  | ATCAACACCTTCATCTCCACCCACACACACGCCA | 900          |
|-----------------------|-------|---|------------------------------------|--------------|
| Qy                    | 901   | GTAGTTATTAAAAA  | AAAAA 920                          |              |
| Db                    | 901   | GTAGTTATTAAAAA  | AAAAA 920                          |              |
| RESULT 9              |       |   |                                    |              |
| HRSRNAG               |       |   |                                    |              |
| LOCUS                 |       |   |                                    |              |
| DEFINITION            |       |   |                                    |              |
| ACCESSION             |       |   |                                    |              |
| VERSION               |       |   |                                    |              |
| KEYWORDS              |       |   |                                    |              |
| SOURCE                |       |   |                                    |              |
| ORGANISM              |       |   |                                    |              |
| REFERENCE             |       |   |                                    |              |
| AUTHORS               |       |   |                                    |              |
| TITLE                 |       |   |                                    |              |
| JOURNAL               |       |   |                                    |              |
| MEDLINE               |       |   |                                    |              |
| PUBMED                |       |   |                                    |              |
| FEATURES              |       |   |                                    |              |
| source                |       |   |                                    |              |
| CDS                   |       |   |                                    |              |
| BASE COUNT            | 379   | a   | 290                                | c            |
| ORIGIN                |       |   |                                    |              |
| Query Match           | 99.1% | Score   | 911.8                              | DB 14        |
| Best Local Similarity | 99.8% | Pred.   | No. 1.9e-190                       |              |
| Matches               | 913   | Conservative  | 0                                  | Mismatches 2 |
|                       |       |   |                                    | Indels 0     |
|                       |       |   |                                    | Gaps 0       |
| Qy                    | 1     | TGCAAAACATGTCCAAAAACAGGACCAACGACCGCTAAGACACTAGAAAGACCTGGGA    | 60                                 |              |
| Db                    | 9     | TGCAAAACATGTCCAAAAACAGGACCAACGACCGCTAAGACACTAGAAAGACCTGGGA    | 68                                 |              |
| Qy                    | 61    | CACCTCTCAATCATTTATTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT | 120                                |              |
| Db                    | 69    | CACCTCTCAATTTATTATTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTAT | 128                                |              |
| Qy                    | 121   | AGCAAAATCATCATTTCCATTCCTGGCAATGATAAATCTCAACTTCATTATAAATACAGC  | 180                                |              |
| Db                    | 129   | AGCAAAATCATCATTTCCATTCCTGGCAATGATAAATCTCAACTTCATTATAAATACAGC  | 188                                |              |
| Qy                    | 181   | CATCATATTCATAGCCTCGGCAACCAACCAAGTCACACTAACATCATCAAGA          | 240                                |              |
| Db                    | 189   | CATCATATTCATAGCCTCGGCAACCAACCAAGTCACACTAACATCATCAAGA          | 248                                |              |
| Qy                    | 241   | TGCAACAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGATCTCTCAGCTTGG    | 300                                |              |
| Db                    | 249   | TGCAACAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGATCTCTCAGCTTGG    | 308                                |              |
| Qy                    | 301   | AATCAGCTTCTCCAATCTGTGTGAATTTATCATCAAAACCAACCACTAGCTTCAAC      | 360                                |              |



```

Source
1. .922
/organism="Human respiratory syncytial virus"
/mol_type="genomic RNA"
/strain="subgroup A"
/isolate="MON 7 91 (Montevideo/Uruguay, 1991)"
/db_xref="taxon:11250"
/cell_line="HEP-2"
/gene="G"
16. .912
16. .912
16. .912
/partial
/gene="G"
/codon_start=1
/product="glycoprotein"
/protein_id="CAA83872.1"
/db_xref="GI:485889"
/db_xref="SPTREMBL:O82068"
/translation="ISKNDQRTAKTLEKWTDLNLYLFTSSGLYKLNLSIAQITLS
ILAMISLTIIITAIPIASANKHVLTITAIQDTSOIKNTTPTLYLTQDPOLGIFS
NLSEITSTQITLSTTGVSKNLQPTTKTKNTTQIOTPSKPTTKQKQKPNKPN
NDFHEVFNFVPCISCSNPTCWAICRIKPKKPKTKTKPTKTKTKKDLKPO
TKPKVEPTTKPTKPTKNTTKNTITLLTNNTGNPKLTSQMETHTSSEGNLSP
SQVSTSEHPSQSPSPNTTQ"
BASE COUNT 378 a 291 c 97 g 156 t
ORIGIN
Query Match 98.8%; Score 909.2; DB 14; Length 922;
Best Local Similarity 99.7%; Pred. No. 7.1e-190;
Matches 911; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TGCACACATGTCACAAACACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60
DB 9 TGCACACATGTCACAAACACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 68
QY 61 CACTCTCAATATTTATTATTATCATATCATCGGCTTATATAGTTAATCTTAATCTGT 120
DB 69 CACTCTCAATATTTATTATTATCATATCATCGGCTTATATAGTTAATCTTAATCTAT 128
QY 121 AGCACAAATCACATTCATCTCGCAATGATATCTCAACTTCACATTACATTATACAGC 180
DB 129 AGCACAAATCACATTCATCTCGCAATGATATCTCAACTTCACATTACATTATACAGC 188
QY 181 CATCATATTATAGCTCGGCAACCAACAAAGTCACTACATCAACCTCAATCATACAGA 240
DB 189 CATCATATTATAGCTCGGCAACCAACAAAGTCACTACATCAACCTCAATCATACAGA 248
QY 241 TGCACACAGCAGATCAGACACACACCCACATACCTGCTCAGATCTCAGCTGG 300
DB 249 TGCACACAGCAGATCAGACACACACCCACATACCTGCTCAGATCTCAGCTGG 308
QY 301 AATCAGCTTCTCCAATCTGCTGGAATTTACATCAACACACACACCATCTAGCTTCAAC 360
DB 309 AATCAGCTTCTCCAATCTGCTGGAATTTACATCAACACACACCATCTAGCTTCAAC 368
QY 361 AACACAGAGTCAAGTCAACCTGCACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
DB 369 AACACAGAGTCAAGTCAACCTGCACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAG 428
QY 421 AACCCAAACACACACCCAGCCCACTACAAACACAGCCCAACACACACACACACAA 480
DB 429 AACCCAAACACACACCCAGCCCACTACAAACACAGCCCAACACACACACACAA 488
QY 481 ACCCAATATGATTTTCACTTCGAAGTGTTTAACTTTGTACCTGCGAGCATATGCACAA 540
DB 489 ACCCAATATGATTTTCACTTCGAAGTGTTTAACTTTGTACCTGCGAGCATATGCACAA 548
QY 541 CAATCCAACTGCTGGCTATCTGCACCAAGATACCAACCAACCAACCAACCAACCA 600
DB 549 CAATCCAACTGCTGGCTATCTGCACCAAGATACCAACCAACCAACCAACCAACCA 608
QY 601 CACCCACAGCTTACAAACCAACCACTTCAAGCAACCAACCAACCAACCAACCA 660
DB 609 CACCCACAGCTTACAAACCAACCACTTCAAGCAACCAACCAACCAACCAACCA 668

```

[illegible]

RESULT 12  
AR080424  
LOCUS AR080424  
DEFINITION Sequence 28 from patent US 5968776.  
ACCESSION AR080424  
VERSION AR080424.1 GI:10007159  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.  
REFERENCE 1 (bases 1 to 894)  
AUTHORS Klein,M.H., Du,R.-P. and Ewasyshyn,M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers

|  |  |  |            |    |        |    |      |    |  |
|--|--|--|------------|----|--------|----|------|----|--|
| protection against parainfluenza virus and respiratory syncytial virus |  |  |            |    |        |    |      |    |  |
| JOURNAL FEATURES   | Patent: US 5968776-A 28 19-OCT-1999; Location/Qualifiers 1. .894 /organism="unknown" |  |            |    |        |    |      |    |  |
| BASE COUNT   | 364  | a  | 288        | c  | 93     | g  | 149  | t  |  |
| ORIGIN   |  |  |            |    |        |    |      |    |  |
| Query Match  | 97.2%; Score 894; DB 6; Length 894;  |  |            |    |        |    |      |    |  |
| Best Local Similarity  | 100.0%; Pred. No. 1.6e-186;  |  |            |    |        |    |      |    |  |
| Matches 894;   | Conservative   | 0;   | Mismatches | 0; | Indels | 0; | Gaps | 0; |  |
| QY   | 8  | ATGTCCAAAACAAAGGACCAAGCAGCAGCTTAAGACACTAGAAAAGACCTGGGACACTTCTC   | 67         |    |        |    |      |    |  |
| Db   | 1  | ATGTCCAAAACAAAGGACCAAGCAGCAGCTTAAGACACTAGAAAAGACCTGGGACACTTCTC   | 60         |    |        |    |      |    |  |
| QY   | 68   | AATCATTTTATTATTTCATATCATCATCGGGCTTATAAAGTTAAATCTTTAAATCTGTAGACAA | 127        |    |        |    |      |    |  |
| Db   | 61   | AATCATTTTATTATTTCATATCATCGGGCTTATAAAGTTAAATCTTTAAATCTGTAGACAA    | 120        |    |        |    |      |    |  |
| QY   | 128  | ATCATATTATTCATTCTTGGCAATGATTAATCTCAACTTCATTATTAATTACAGCCATCATTA  | 187        |    |        |    |      |    |  |
| Db   | 121  | ATCATATTATTCATTCTTGGCAATGATTAATCTCAACTTCATTATTAATTACAGCCATCATTA  | 180        |    |        |    |      |    |  |
| QY   | 188  | TTCATAGCTTCGGCAAAACCAAGTTCAGTAACTAACAATCTGAATCATACAAGTATGCAACA   | 247        |    |        |    |      |    |  |
| Db   | 191  | TTCATAGCTTCGGCAAAACCAAGTTCAGTAACTAACAATCTGAATCATACAAGTATGCAACA   | 240        |    |        |    |      |    |  |
| QY   | 248  | AGCCAGATCAAGAACACAAACCCCAACATCCTCACTCAGGATCCTCAGCTTTGGAATCAGC    | 307        |    |        |    |      |    |  |
| Db   | 241  | AGCCAGATCAAGAACACAAACCCCAACATCCTCACTCAGGATCCTCAGCTTTGGAATCAGC    | 300        |    |        |    |      |    |  |
| QY   | 308  | TTCTCCAAATCTGTCTGAAATTTATCATCAAAACCAACCACTATAGCTTCAACACACCA      | 367        |    |        |    |      |    |  |
| Db   | 301  | TTCTCCAAATCTGTCTGAAATTTATCATCAAAACCAACCACTATAGCTTCAACACACCA      | 360        |    |        |    |      |    |  |
| QY   | 368  | GGAGTCAAGTCMAACTGCAACCCACCAACAGTCAAGCTTAAAGCTTAAACACACACCCAA     | 427        |    |        |    |      |    |  |
| Db   | 361  | GGAGTCAAGTCMAACTGCAACCCACCAACAGTCAAGCTTAAAGCTTAAACACACACCCAA     | 420        |    |        |    |      |    |  |
| QY   | 428  | ACACAAACCCAGCAGCCCACTACAAAACACGCGCAAAACCAACCACTATAGCTTCAACACACCA | 487        |    |        |    |      |    |  |
| Db   | 421  | ACACAAACCCAGCAGCCCACTACAAAACACGCGCAAAACCAACCACTATAGCTTCAACACCA   | 480        |    |        |    |      |    |  |
| QY   | 488  | AATGATTTTTCATCTCGAAGTGTTTAACTTTGTTACCTTGAGCATATGCGACAACTTCCA     | 547        |    |        |    |      |    |  |
| Db   | 481  | AATGATTTTTCATCTCGAAGTGTTTAACTTTGTTACCTTGAGCATATGCGACAACTTCCA     | 540        |    |        |    |      |    |  |
| QY   | 548  | ACCTGCTGGGCTATCTCAAAAGAAATACCAAAACAAAACCCAGGAAGGAAACCAACCC       | 607        |    |        |    |      |    |  |
| Db   | 541  | ACCTGCTGGGCTATCTCAAAAGAAATACCAAAACAAAACCCAGGAAGGAAACCAACCC       | 600        |    |        |    |      |    |  |
| QY   | 608  | AAGCCTTCAAAAAACCAACCTTTCAAGACAAACCAAAAAAGATCTCAAACCTCAAACCACT    | 667        |    |        |    |      |    |  |
| Db   | 601  | AAGCCTTCAAAAAACCAACCTTTCAAGACAAACCAAAAAAGATCTCAAACCTCAAACCACT    | 660        |    |        |    |      |    |  |
| QY   | 668  | AAACCAAGGAAGTATCCACACCAAGCCCAAGAGAGCCCAACCAATCAACACCAACCAAA      | 727        |    |        |    |      |    |  |
| Db   | 661  | AAACCAAGGAAGTATCCACACCAAGCCCAAGAGAGCCCAACCAATCAACACCAACCAAA      | 720        |    |        |    |      |    |  |
| QY   | 728  | ACAAACATCAAACTACACTGCTCACCACACACACAGGAATTCAAAATCTCAAGT           | 787        |    |        |    |      |    |  |
| Db   | 721  | ACAAACATCAAACTACACTGCTCACCACACACAGGAATTCAAAATCTCAAGT             | 780        |    |        |    |      |    |  |
| QY   | 788  | CAAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTTAAGCCCTTCTCAAGTCTCC      | 847        |    |        |    |      |    |  |
| Db   | 781  | CAAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTTAAGCCCTTCTCAAGTCTCC      | 840        |    |        |    |      |    |  |
| QY   | 848  | ACAAATCCGAGCAGCCATCAAAACCTCATCTCGACCCAAACAAACAGCCAG              | 901        |    |        |    |      |    |  |
| Db   | 841  | ACAAATCCGAGCAGCCATCAAAACCTCATCTCGACCCAAACAAACAGCCAG              | 894        |    |        |    |      |    |  |

## RESULT 13

AR092548  
LOCUS AR092548 894 bp DNA linear PAT 08-SEP-2000  
DEFINITION Sequence 28 from patent US 5998169.  
ACCESSION AR092548  
VERSION AR092548.1 GI:10019302  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 894)  
AUTHORS Klein,M.H., Du,R.-P. and Ewaszyzyn,M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus

JOURNAL Patent: US 5998169-A 28 07-DEC-1999;  
FEATURES Location/Qualifiers  
source 1..894  
BASE COUNT 364 a 288 c 93 g 149 t  
ORIGIN  
Query Match 97.2%; Score 894; DB 6; Length 894;  
Best Local Similarity 100.0%; Pred. No. 1.6e-186;  
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ATGTCCAAAACAAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGACACTCTC 67  
DB 1 ATGTCCAAAACAAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGACACTCTC 60  
QY 68 AATCATTATTATTTCATATCATCTGGCTTATATTAAGTAAATCTTAAATCTGTAGCAAA 127  
DB 61 AATCATTATTATTTCATATCATCTGGCTTATATTAAGTAAATCTTAAATCTGTAGCAAA 120  
QY 128 ATCACTATTATCCATCTTGGCAATGATATCTCACTTCACTTATATTAATACAGGCATCAT 187  
DB 121 ATCACTATTATCCATCTTGGCAATGATATCTCACTTCACTTATATTAATACAGGCATCAT 180  
QY 188 TTCTAGCTCGGCAAAACCAAGTCACTAACAATGCAATCATACAAAGTGCACAA 247  
DB 181 TTCTAGCTCGGCAAAACCAAGTCACTAACAATGCAATCATACAAAGTGCACAA 240  
QY 248 AGCCAGATCAAGACACAAACCCCACTACTCTCAGGATCTCTAGTTGGATCAGC 307  
DB 241 AGCCAGATCAAGACACAAACCCCACTACTCTCAGGATCTCTAGTTGGATCAGC 300  
QY 308 TTCTCCAATCTGTCTGAATTTATCATCAAAACCAACCACTAGCTTCAACCAACCA 367  
DB 301 TTCTCCAATCTGTCTGAATTTATCATCAAAACCAACCACTAGCTTCAACCAACCA 360  
QY 368 GGAGTCAAGTCAAACTTGGCAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 427  
DB 361 GGAGTCAAGTCAAACTTGGCAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 420  
QY 428 ACACACCCAGCAGCCACTTACAAACACGCGCAACCAACCAACCAACCAACCAACCAAT 487  
DB 421 ACACACCCAGCAGCCACTTACAAACACGCGCAACCAACCAACCAACCAACCAACCAAT 480  
QY 488 AATGATTTCCTTCAAGTGTGTTAACTTTGTACCTCTGAGCATATGACCAACATCCA 547  
DB 481 AATGATTTCCTTCAAGTGTGTTAACTTTGTACCTCTGAGCATATGACCAACATCCA 540  
QY 548 ACTGTCTGGGTATCTCAAAAAGATACCAACCAAAAACCGAGAAAGAAACCAACACC 607  
DB 541 ACTGTCTGGGTATCTCAAAAAGATACCAACCAAAAACCGAGAAAGAAACCAACACC 600  
QY 608 AAGCTTCAAAAACCAACCTTCAAGACACCAACCAAAAAGATCTCAAACTCAAACT 667  
DB 601 AAGCTTCAAAAACCAACCTTCAAGACACCAACCAAAAAGATCTCAAACTCAAACT 660  
QY 668 AAACCAAGGAAGTACCCACCAACCAAGCCCAAGAGAGCAACCATCAACACCAACAA 727

Db 661 AAACCAAGGAAGTACCCACCAAGCCCAAGAGCCCAAGACCAACCATCAACACCAAA 720  
QY 728 ACAAACATCAAACTACTGTCTCAACAAACACCAAGGAATCCAAATCTCAAACT 787  
Db 721 ACAAACATCAAACTACTGTCTCAACAAACACCAAGGAATCCAAATCTCAAACT 780  
QY 788 CAATGGAAACCTTCCACTCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 847  
Db 781 CAATGGAAACCTTCCACTCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 840  
QY 848 ACAAACATCGAGCACCCATCAACAAACCTCTCTCAACCAACCAACCAACCAACGCGCAG 901  
Db 841 ACAAACATCGAGCACCCATCAACAAACCTCTCTCAACCAACCAACCAACCAACGCGCAG 894

RESULT 14  
RSHICE 8510 bp ss-RNA linear VRL 29-NOV-2000  
LOCUS Human respiratory syncytial virus nonstructural protein (1C),  
DEFINITION nonstructural protein (1B), major nucleocapsid (N), phosphoprotein  
(P), protein (M), 1A (1A), G (G), protein (F) and  
envelope-associated protein (22K) gene, complete cds.  
ACCESSION M11486 K01459 K03348 K03349 M11217 M11244 M11487 M11505  
VERSION M11514 M11631 M12966  
KEYWORDS M11486.1 GI:333925  
envelope-associated protein; fusion glycoprotein; major  
nucleocapsid protein; major surface glycoprotein; matrix protein;  
nonstructural protein; phosphoprotein.  
SOURCE Human respiratory syncytial virus  
ORGANISM Human respiratory syncytial virus  
Viruses; ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE 1 (bases 1085 to 2220; 5254 to 5544)  
AUTHORS Elango,N. and Venkatesan,S.  
TITLE Amino acid sequence of human respiratory syncytial virus  
nucleocapsid protein  
JOURNAL Nucleic Acids Res. 11 (17), 5941-5951 (1983)  
MEDLINE 83299261  
PUBMED 6310521  
REFERENCE 2 (bases 3211 to 4157)  
AUTHORS Satake,M. and Venkatesan,S.  
TITLE Nucleotide sequence of the gene encoding respiratory syncytial  
virus matrix protein  
JOURNAL J. Virol. 50 (1), 92-99 (1984)  
MEDLINE 84138836  
PUBMED 6699948  
REFERENCE 3 (bases 2288 to 3191)  
AUTHORS Satake,M., Elango,N. and Venkatesan,S.  
TITLE Sequence analysis of the respiratory syncytial virus phosphoprotein  
Gene  
JOURNAL J. Virol. 52 (3), 991-994 (1984)  
MEDLINE 85033973  
PUBMED 6548527  
REFERENCE 4 (bases 5602 to 7500)  
AUTHORS Collins,P.L., Huang,Y.T. and Wertz,G.W.  
TITLE Nucleotide sequence of the gene encoding the fusion (F)  
glycoprotein of human respiratory syncytial virus  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (24), 7683-7687 (1984)  
MEDLINE 85088471  
PUBMED 6096849  
REFERENCE 5 (bases 7551 to 8510)  
AUTHORS Collins,P.L. and Wertz,G.W.  
TITLE The envelope-associated 22K protein of human respiratory syncytial  
virus: nucleotide sequence of the mRNA and a related polytranscript  
JOURNAL J. Virol. 54 (1), 65-71 (1985)  
MEDLINE 85135082  
PUBMED 3838351  
REFERENCE 6 (bases 8 to 1050; 7554 to 8506)  
AUTHORS Elango,N., Satake,M. and Venkatesan,S.  
TITLE mRNA sequence of three respiratory syncytial virus genes encoding  
two nonstructural proteins and a 22K structural protein  
JOURNAL J. Virol. 55 (1), 101-110 (1985)  
MEDLINE 85237684

|           |         |   |   |
|-----------|---------|---|---|
| 4009789   | PUBMED  | 7 (bases 5602 to 7423)  | Elango,N., Satake,M., Coligan,J.E., Norrby,E., Camargo,E. and Venkatesan,S. |
| REFERENCE | AUTHORS |   |   |
| TITLE     |         | Respiratory syncytial virus fusion glycoprotein: nucleotide sequence of mRNA, identification of cleavage activation site and amino acid sequence of N-terminus of F1 subunit  |   |
| JOURNAL   |         | Nucleic Acids Res. 13 (5), 1559-1574 (1985)   |   |
| MEDLINE   |         | 85215565  |   |
| PUBMED    |         | 2987829   |   |
| REFERENCE |         | 8 (bases 4630 to 5543)  |   |
| AUTHORS   |         | Satake,M., Coligan,J.E., Elango,N., Norrby,E. and Venkatesan,S.   |   |
| TITLE     |         | Respiratory syncytial virus envelope glycoprotein (G) has a novel structure   |   |
| JOURNAL   |         | Nucleic Acids Res. 13 (21), 7795-7812 (1985)  |   |
| MEDLINE   |         | 86067198  |   |
| PUBMED    |         | 4069997   |   |
| REFERENCE |         | 9 (bases 4627 to 5544)  |   |
| AUTHORS   |         | Wertz,G.W., Collins,P.L., Huang,Y., Gruber,C., Levine,S. and Ball,L.A.  |   |
| TITLE     |         | Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein  |   |
| JOURNAL   |         | Proc. Natl. Acad. Sci. U.S.A. 82 (12), 4075-4079 (1985)   |   |
| MEDLINE   |         | 85216636  |   |
| PUBMED    |         | 3858865   |   |
| REFERENCE |         | 10 (bases 4173 to 4571)   |   |
| AUTHORS   |         | Collins,P.L. and Wertz,G.W.   |   |
| TITLE     |         | The 1A protein gene of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polycistronic transcript  |   |
| JOURNAL   |         | Virology 141 (2), 283-291 (1985)  |   |
| MEDLINE   |         | 86098645  |   |
| PUBMED    |         | 3879976   |   |
| REFERENCE |         | 11 (bases 1 to 528; 552 to 1050)  |   |
| AUTHORS   |         | Collins,P.L. and Wertz,G.W.   |   |
| TITLE     |         | Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus  |   |
| JOURNAL   |         | Virology 143 (2), 442-451 (1985)  |   |
| MEDLINE   |         | 86045905  |   |
| PUBMED    |         | 2998021   |   |
| REFERENCE |         | 12 (bases 1081 to 2277)   |   |
| AUTHORS   |         | Collins,P.L., Anderson,K., Langer,S.J. and Wertz,G.W.   |   |
| TITLE     |         | Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus   |   |
| JOURNAL   |         | Virology 146 (1), 69-77 (1985)  |   |
| MEDLINE   |         | 85301974  |   |
| PUBMED    |         | 3839952   |   |
| REFERENCE |         | 13 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578)   |   |
| AUTHORS   |         | Collins,P.L., Dickens,L.E., Buckler-White,A., Olmsted,R.A., Spriggs,M.K., Camargo,E. and Coelingh,K.V.  |   |
| TITLE     |         | Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order   |   |
| JOURNAL   |         | Proc. Natl. Acad. Sci. U.S.A. 83 (13), 4594-4598 (1986)   |   |
| MEDLINE   |         | 86259643  |   |
| PUBMED    |         | 3460060   |   |
| COMMENT   |         | Reprint for [2] and clean copy sequence for [12], [5] kindly provided by P.Collins, 21-FEB-1986.<br>Respiratory syncytial virus is a negative-strand RNA pleomorphic enveloped virus of the genus Pneumovirus. The matrix protein has no homology with the matrix proteins of other negative-stranded RNA viruses, implying that RS virus has undergone extensive evolutionary divergence. Two unidentified reading frames which potentially encode proteins were located: one overlaps the matrix protein and the other the F protein. The biological significance of these two reading frames is not clear. The positive strand is shown. |   |
| FEATURES  | source  | Location/Qualifiers<br>1..8510<br>/organism="Human respiratory syncytial virus"<br>/mol_type="genomic RNA"<br>/strain="A2"  |   |

Query Match

92.8%; Score 853.4; DB 14; Length 8510;

```
/db_xref="taxon:11250"
/lab_host="HEP-2 cell"
1..528
/gene="1C"
1..528
/gene="1C"
55..474
/gene="1C"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAB59850.1"
/db_xref="GI:333926"
/translation="MGSNSLSMIKVRQLQMLFNDVEALLKTCYTDKLIHLTNALAKA
VLIHTKLNGIVFVHVTISSDICPNNIVKSNFTMPVLQNGGYIWEWMLTHCSQPN
GLLDNDCEIKFSKSLSDSTMTNMQLSGLLFDLNP"
171
/gene="1C"
/replace="t"
552..1050
/gene="1B"
552..1050
/gene="1B"
584..958
/gene="1B"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAB59851.1"
/db_xref="GI:333927"
/translation="MDTTHNDNTPORLMTDMRPLSLETTITSLTRDITTHKFIYLI
HECIVRKDEKQATFELVNYEMKLLHKVGSYKKYKYTEYNTKYTFPMPPIFNDGP
LECGIKTKTKPIIYKYDLNP"
747
/gene="1B"
/replace="g"
769
/gene="1B"
/replace="a"
1081..2277
/gene="N"
1081..2277
/gene="N"
1096..2271
/gene="N"
/codon_start=1
/product="major nucleocapsid"
/protein_id="AAB59852.1"
/db_xref="GI:333928"
/translation="MALSQVKNLNTLNKQLSSSKYTIQRTSGDSIDTPNYDVQKHI
NKLQMLLITEDANHKFTGLIGLYAMSLRGREDTIKLRDAGYHVKANGVDVTHHQ
DINGKMEFVLTSLTTEIQINIEISRSYKQMLKMGVAPYERHDSDCQCI
LCIALVTTLKAAGDRSLTAVIRRANVLNEMKRYKGLLPKOLANSFYEVFKHPH
FIDVPHFGIAAQSTGRGSRVEGIFAGLPMNAYGQVMLNGVLAKSVKIMLGHAS
VQAMEQVVEYVYAKLQGEAGFYHILNPKASLLSTQPFHFSVVILGNAAGIGIM
GEYRGTPRNQDLYDAKAYASQLKENGVINYSVLDTABELEATKHQLNPKNDQVEL"
2289..3196
/gene="p"
2289..3196
/gene="p"
2302..3027
/gene="p"
/codon_start=1
/product="phosphoprotein"
/protein_id="AAB59853.1"
/db_xref="GI:333929"
/translation="MEKPAFEHGEDANNRATKFLSIKGFSTKDPKPKKDSIISVN
SIDIEVTKBSPTINSTINETDNTAGNPNYQKPLVSFKEDPTSPNSPKLY
KETIETFDNNEESSYSYERINDQNDNTIARLDRIEKLGLMLHTLVVASAGPT
SARDGIRDAMIGLREEMIEKIRTEALMTNDRLEAMRNEESEKWKADTSDVSLNP
```

Best Local Similarity 95.5%; Pred. No. 1.2e-177; Mismatches 41; Indels 0; Gaps 0; Matches 876; Conservative 0;

Qy 1 TGCACCATGTCACAAACCAAGGACCAAGCAGCCGCTAAGACACTAGAAAGACCTGGGA 60  
Db 4634 TGCACCATGTCACAAACCAAGGACCAAGCAGCCGCTAAGACACTAGAAAGACCTGGGA 4693

Qy 61 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATTAAGTTAAATCTTAATCTGT 120  
Db 4694 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATTAAGTTAAATCTTAATCTGT 4753

Qy 121 AGCACAATCATATTCATCTGCGCAATGATATATCTCAACTTCACCTTATATTAATTCAGC 180  
Db 4754 AGCACAATCATATTCATCTGCGCAATGATATATCTCAACTTCACCTTATATTAATTCAGC 4813

Qy 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACTCAACACTGCAATCATACAGA 240  
Db 4814 CATCATATTCATAGCTCGGCAACCAACCAAGTCACTCAACACTGCAATCATACAGA 4873

Qy 241 TGCACCAAGCCAGATCAAGAACACACCCCAACATACCTGCTCAGAGATCTCAGCTGG 300  
Db 4874 TGCACCAAGCCAGATCAAGAACACACCCCAACATACCTGCTCAGAGATCTCAGCTGG 4933

Qy 301 AATCAGCTTCTCCTCAATCTGCTGAAATTTACATCAACACCAACCACTACTAGCTTCAAC 360  
Db 4934 AATCAGCTTCTCCTCAATCTGCTGAAATTTACATCAACACCAACCACTACTAGCTTCAAC 4993

Qy 361 AACCAAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAG 420  
Db 4994 AACCAAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAG 5053

Qy 421 AACCCAAACACACCCAGCAGCCCACTCAAAACCAAGCAGCCCAACCAACCAACCAACCA 480  
Db 5054 AACTCAAAACACACCCAGCAGCCCACTCAAAACCAAGCAGCCCAACCAACCAACCAACCA 5113

Qy 481 ACCCAATATGATTTCTGAGTGTTCATCTGAGTGTTCATCTGAGTGTTCATCTGAGTGTTC 540  
Db 5114 ACCCAATATGATTTCTGAGTGTTCATCTGAGTGTTCATCTGAGTGTTCATCTGAGTGTTC 5173

Qy 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGGAGGAGGAGGAGG 600  
Db 5174 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGGAGGAGGAGGAGG 5233

Qy 601 CACCAACCAAGCTACAAAACCAACCACTTCAAGACCAACCAAAAGATCTCAACCTCA 660  
Db 5234 CACTACCAAGCTTCAAAAACCAACCACTTCAAGACCAACCAAAAGATCTCAACCTCA 5293

Qy 661 AACCACTAAACCAAGGAGTACCCCAACCAAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAG 720  
Db 5294 AACCACTAAATCAAGGAGTACCCCAACCAAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAG 5353

Qy 721 CACCAAAACAAACATCAACACTACTGCTCACCACCAACCAACCAAGGAGGAGGAGGAGGAG 780  
Db 5354 CACCAAAACAAACATCAACACTACTGCTCACCACCAACCAACCAAGGAGGAGGAGGAGGAG 5413

Qy 781 CACAAGTCAATGGAACCTTCACTCAACTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
Db 5414 CACAAGTCAATGGAACCTTCACTCAACTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5473

Qy 841 AGTCTCCCAACATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
Db 5474 AGTCTCTAACATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5533

Qy 901 GTAGTTATTAATAAAAAA 919  
Db 5534 GTAGTTACTTAATAACATA 5552

RESULT 15  
AR093219  
LOCUS AR093219 15222 bp DNA linear PAT 08-SEP-2000  
DEFINITION Sequence 23 from patent US 5998602.  
ACCESSION AR093219

VERSION AR093219.1 GI:10019970  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 15222)  
AUTHORS Torrence,P.F., Silverman,R.Hugh., Cirino,N.Mario., Li,G. and Xiao,W.  
TITLE RNase L activators and antisense oligonucleotides effective to treat RSV infections  
JOURNAL Patent: US 5998602-A 23 07-DEC-1999;  
FEATURES Location/Qualifiers  
source 1..15222 /organism="unknown"  
BASE COUNT 5924 a 2704 c 2356 g 4238 t  
ORIGIN  
Query Match 92.8%; Score 853.4; DB 6; Length 15222;  
Best Local Similarity 95.5%; Pred. No. 1.2e-177; Mismatches 41; Indels 0; Gaps 0; Matches 876; Conservative 0;

Qy 1 TGCACCATGTCACAAACCAAGGACCAAGCAGCCGCTAAGACACTAGAAAGACCTGGGA 60  
Db 4681 TGCACCATGTCACAAACCAAGGACCAAGCAGCCGCTAAGACACTAGAAAGACCTGGGA 4740

Qy 61 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATTAAGTTAAATCTTAATCTGT 120  
Db 4741 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATTAAGTTAAATCTTAATCTGT 4800

Qy 121 AGCACAATCATATTCATCTGCGCAATGATATCTCAACTTCACCTTATATTAATTCAGC 180  
Db 4801 AGCACAATCATATTCATCTGCGCAATGATATCTCAACTTCACCTTATATTAATTCAGC 4860

Qy 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACTCAACACTGCAATCATACAGA 240  
Db 4861 CATCATATTCATAGCTCGGCAACCAACCAAGTCACTCAACACTGCAATCATACAGA 4920

Qy 241 TGCACCAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300  
Db 4921 TGCACCAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 4980

Qy 301 AATCAGCTTCTCCTCAATCTGCTGAAATTTACATCAACCAACCAAGGAGGAGGAGGAGG 360  
Db 4981 AATCAGCTTCTCCTCAATCTGCTGAAATTTACATCAACCAACCAAGGAGGAGGAGGAGG 5040

Qy 361 AACCAAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAG 420  
Db 5041 AACCAAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAG 5100

Qy 421 AACCCAAACCAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
Db 5101 AACTCAAAACCAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5160

Qy 481 ACCCAATATGATTTCTGAGTGTTCATCTGAGTGTTCATCTGAGTGTTCATCTGAGTGTTC 540  
Db 5161 ACCCAATATGATTTCTGAGTGTTCATCTGAGTGTTCATCTGAGTGTTCATCTGAGTGTTC 5220

Qy 541 CAATCCCAACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGGAGGAGGAGGAG 600  
Db 5221 CAATCCCAACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGGAGGAGGAGGAG 5280

Qy 601 CACCAACCAAGCTTCAAAAACCAACCTTCAAGAACCAACCAAAAGATCTCAAACTCA 660  
Db 5281 CACTTACCAAGCCCAACCAAAACCAACCTTCAAGAACCAACCAAAAGATCTCAAACTCA 5340

Qy 661 AACCACTAAACCAAGGAGTACCCCAACCAAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAG 720  
Db 5341 AACCACTAAATCAAGGAGTACCCCAACCAAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAG 5400

Qy 721 CACCAAAACAAACATCAACACTACTGCTCACCACCAACCAACCAAGGAGGAGGAGGAGGAG 780  
Db 5401 CACCAAAACAAACATCAACACTACTGCTCACCACCAACCAACCAAGGAGGAGGAGGAGGAG 5460

|    |      |   |      |
|----|------|---|------|
| Qy | 781  | CACAAGTCAAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA | 840  |
| Db | 5461 | CACAAGTCAAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA | 5520 |
| Qy | 841  | AGTCTCCACACATCCGAGCACCCTATCAGACCTCTCTCCACCCCAACACACGCCA   | 900  |
| Db | 5521 | AGTCTCTACACATCCGAGTACCCTATCAGACCTCTCTCCACCCCAACACACGCCA   | 5580 |
| Qy | 901  | GTAGTTATTAAAAAAA  | 919  |
| Db | 5581 | GTAGTTACTTAAAAACATA                                       | 5599 |

Search completed: October 30, 2003, 00:01:07  
Job time : 3766.15 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:32:25 ; Search time 70.283 Seconds  
(without alignments)  
1094.143 Million cell updates/sec

Title: US-09-462-816-2  
Perfect score: 1544  
Sequence: 1 MSKNKQRTAKTLEKTWDTL.....VSTTSEHPSPSSPPNTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 1537  | 99.5        | 298    | 12    | Q01929 human respi |
| 2          | 1533  | 99.3        | 298    | 12    | Q82068 human respi |
| 3          | 1393  | 90.2        | 298    | 12    | O09634 human respi |
| 4          | 1389  | 90.0        | 298    | 12    | O09719 respiratory |
| 5          | 1365  | 88.4        | 297    | 12    | Q9YVB5 human respi |
| 6          | 1364  | 88.3        | 297    | 12    | Q82066 human respi |
| 7          | 1356  | 87.8        | 298    | 12    | Q9YVB4 human respi |
| 8          | 1354  | 87.7        | 297    | 12    | Q82071 human respi |
| 9          | 1345  | 87.1        | 297    | 12    | Q82067 human respi |
| 10         | 1344  | 87.0        | 297    | 12    | Q82058 human respi |
| 11         | 1339  | 86.7        | 297    | 12    | O91947 human respi |
| 12         | 1336  | 86.5        | 293    | 12    | Q9YVB3 human respi |
| 13         | 1334  | 86.4        | 295    | 12    | Q86359 respiratory |
| 14         | 1334  | 86.4        | 295    | 12    | Q86360 respiratory |
| 15         | 1334  | 86.4        | 298    | 12    | Q82065 human respi |
| 16         | 1331  | 86.2        | 297    | 12    | Q82074 human respi |

|    |      |      |     |    |        |                    |
|----|------|------|-----|----|--------|--------------------|
| 17 | 1331 | 86.2 | 297 | 12 | Q82064 | Q82064 human respi |
| 18 | 1328 | 86.0 | 297 | 12 | Q82063 | Q82063 human respi |
| 19 | 1326 | 85.9 | 292 | 12 | Q9YVB2 | Q9YVB2 human respi |
| 20 | 1326 | 85.9 | 298 | 12 | Q82077 | Q82077 human respi |
| 21 | 1325 | 85.8 | 298 | 12 | Q82060 | Q82060 human respi |
| 22 | 1324 | 85.8 | 297 | 12 | O91946 | O91946 human respi |
| 23 | 1320 | 85.5 | 297 | 12 | Q82070 | Q82070 human respi |
| 24 | 1319 | 85.4 | 293 | 12 | Q9YVB1 | Q9YVB1 human respi |
| 25 | 1319 | 85.4 | 295 | 12 | Q86361 | Q86361 respiratory |
| 26 | 1319 | 85.4 | 297 | 12 | Q82079 | Q82079 human respi |
| 27 | 1319 | 85.4 | 297 | 12 | Q82057 | Q82057 human respi |
| 28 | 1318 | 85.4 | 297 | 12 | Q9YVC8 | Q9YVC8 human respi |
| 29 | 1317 | 85.3 | 298 | 12 | Q82075 | Q82075 human respi |
| 30 | 1317 | 85.3 | 298 | 12 | Q82062 | Q82062 human respi |
| 31 | 1317 | 85.3 | 298 | 12 | Q82056 | Q82056 human respi |
| 32 | 1316 | 85.2 | 297 | 12 | Q82072 | Q82072 human respi |
| 33 | 1315 | 85.2 | 295 | 12 | Q86357 | Q86357 respiratory |
| 34 | 1314 | 85.1 | 297 | 12 | O91944 | O91944 human respi |
| 35 | 1314 | 85.1 | 297 | 12 | Q82078 | Q82078 human respi |
| 36 | 1314 | 85.1 | 298 | 12 | Q82073 | Q82073 human respi |
| 37 | 1314 | 85.1 | 298 | 12 | Q82076 | Q82076 human respi |
| 38 | 1312 | 85.0 | 298 | 12 | O11978 | O11978 human respi |
| 39 | 1311 | 84.9 | 292 | 12 | Q9YVB0 | Q9YVB0 human respi |
| 40 | 1311 | 84.9 | 298 | 12 | O91945 | O91945 human respi |
| 41 | 1308 | 84.7 | 296 | 12 | Q86358 | Q86358 respiratory |
| 42 | 1306 | 84.6 | 298 | 12 | O8JQU4 | O8JQU4 human respi |
| 43 | 1306 | 84.6 | 298 | 12 | Q82069 | Q82069 human respi |
| 44 | 1305 | 84.5 | 298 | 12 | Q82061 | Q82061 human respi |
| 45 | 1298 | 84.1 | 295 | 12 | Q86356 | Q86356 respiratory |

ALIGNMENTS

RESULT 1

Q01929 ID Q01929 PRELIMINARY; PRT; 298 AA.  
AC Q01929;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE G protein.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91065351; PubMed=2249671;  
RA Garcia-Barrero B., Portela A., Delgado T., Lopez J.A., Melero J.A.;  
RT "Frame shift mutations as a novel mechanism for the generation of  
RT newralization resistant mutants of human respiratory syncytial  
RT virus.";  
RL EMBO J. 9:4181-4187(1990).  
DR EMBL; X17085; CAA34937.1; -;  
DR InterPro; IPR000925; Glycoprot G  
DR InterPro; IPR006162; Ppantne attach.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.  
SQ SEQUENCE 298 AA; 32771 MW; F8CD4213D97C2952 CRC64;

Query Match 99.5%; Score 1537; DB 12; Length 298;  
Best Local Similarity 99.3%; Pred. No. 1e-105;  
Matches 296; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

|    |    |   |     |
|----|----|---|-----|
| Qy | 1  | MSKNKQRTAKTLEKTWDTLNLHLLFTSSGLYKLNLSVAQITLSILAMITSLITAI   | 60  |
| Db | 1  | MSKNKQRTAKTLEKTWDTLNLHLLFTSSGLYKLNLSVAQITLSILAMITSLITAI   | 60  |
| Qy | 61 | FIASANKVLTITAIQDATSQIKNTPTTYLTQDPQLGHSFNSLSEITSTTTILASTTP | 120 |
| Db | 61 | FIASANKVLTITAIQDATSQIKNTPTTYLTQDPQLGHSFNSLSEITSTTTILASTTP | 120 |



```
QY 121 GVSNNLOPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
Db 121 GVSNNLOPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
QY 181 TCWAIKCRIPNKKPGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTEPTINTTK 240
Db 181 TCWAIKCRIPNKKPGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTEPTINTTK 240
QY 241 TMTTLLTNNTGNPKLTSMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298
Db 241 TMTTLLTNNTGNPKLTSMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298

RESULT 2
Q82068 PRELIMINARY; PRT; 298 AA.
AC Q82068;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrasile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein."
RL J. Virol. 68:5448-5459(1994).
RL EMBL; Z33429; CAA83872.1; -
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Pplantne_attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
FT NON TER 1
SQ SEQUENCE 298 AA; 32753 MW; 7D080307897A772B CRC64;

Query Match 99.3%; Score 1533; DB 12; Length 298;
Best Local Similarity 99.0%; Pred. No. 2e-105;
Matches 295; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTLIITAI 60
Db 1 ISKNDQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTLIITAI 60
QY 61 FIASANKVLTTLAIQDATSQIKNTPTVLTQDPQLGIFSFLSLSITSTTILASTTP 120
Db 61 FIASANKVLTTLAIQDATSQIKNTPTVLTQDPQLGIFSFLSLSITSTTILASTTP 120
QY 121 GVSNNLOPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
Db 121 GVSNNLOPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
QY 181 TCWAIKCRIPNKKPGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTEPTINTTK 240
Db 181 TCWAIKCRIPNKKPGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTEPTINTTK 240
QY 241 TMTTLLTNNTGNPKLTSMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298
Db 241 TMTTLLTNNTGNPKLTSMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298

RESULT 3
O09634 PRELIMINARY; PRT; 298 AA.
ID O09634
```

```
AC O09634;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Attachment protein (G).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RX MEDLINE=97185152; PubMed=9032893;
RA Tolley K.P., Marriott A.C., Simpson A., Flows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle C.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL Vaccine 14:1637-1646(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RA Easton A.J.;
RL Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39662; AAC57026.1; -
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Pplantne_attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
SQ SEQUENCE 298 AA; 32749 MW; 60F27B29D4997F31 CRC64;

Query Match 90.2%; Score 1393; DB 12; Length 298;
Best Local Similarity 91.3%; Pred. No. 4.3e-95;
Matches 272; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTLIITAI 60
Db 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTLIITAI 60
QY 61 FIASANKVLTTLAIQDATSQIKNTPTVLTQDPQLGIFSFLSLSITSTTILASTTP 120
Db 61 FIASANKVLTTLAIQDATSQIKNTPTVLTQDPQLGIFSFLSLSITSTTILASTTP 120
QY 121 GVSNNLOPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
Db 121 GVSNNLOPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
QY 181 TCWAIKCRIPNKKPGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTEPTINTTK 240
Db 181 TCWAIKCRIPNKKPGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTEPTINTTK 240
QY 241 TMTTLLTNNTGNPKLTSMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298
Db 241 TMTTLLTNNTGNPKLTSMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298

RESULT 4
O09719 PRELIMINARY; PRT; 298 AA.
ID O09719
AC O09719;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Attachment glycoprotein (G).
GN G.
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2 ts1C;
```

```

RX MEDLINE=97185152; PubMed=9032893;
RA Tolley K.P., Marriott A.C., Simpson A., Plows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle C.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL vaccine 14:1637-1646 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S2 ts1c;
RA Easton A.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39661; AAC57036.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
DR SEQUENCE 298 AA; 32775 MW; 67FA043682FA450 CRC64;

Query Match      90.0%; Score 1389; DB 12; Length 298;
Best Local Similarity 90.9%; Pred. No. 8.4e-95;
Matches 271; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANKHVLTTLAIQDATSQIKNTTPTYLTPQDGLGSPSNLSEITTSQTTILASTTP 120
DB 61 FIASANKHVLTTLAIQDATSQIKNTTPTYLTPQDGLGSPSNLSEITTSQTTILASTTP 120

QY 121 GVKSNLQPTTVKTKNTTTTQTPSKPTTKQKQKPPKPNNDFFHFEVFNVPVCSICSNP 180
DB 121 SVKSTLQSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNNDFFHFEVFNVPVCSICSNP 180

QY 181 TCWAIKRIKPKKPGKKTITTKPTKPKTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240
DB 181 TCWAIKRIKPKKPGKKTITTKPTKPKTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240

QY 241 TNIITTLTNTTGNPKLTSQMETFHSSTSGNLSPSQVSTTSEHPSPSPNTTQ 298
DB 241 TNIITTLTNTTGNPKLTSQMETFHSSTSGNLSPSQVSTTSEHPSPSPNTTQ 298

RESULT 5
QYVVB5 PRELIMINARY; PRT; 297 AA.
AC QYVVB5
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Attachment glycoprotein G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HV2780;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Murson M.A., Prince G., Anderson L.J., Wertz G.W.;
RA "Antigenic and genetic diversity among the attachment proteins of
RA group A respiratory syncytial viruses that have caused repeat
RA infections in children.";
RL J. Infect. Dis. 178:925-932 (1998).
DR EMBL; AF065405; AAD02941.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
DR SEQUENCE 297 AA; 32572 MW; 5C10E719A05EF5C1 CRC64;

Query Match      88.4%; Score 1365; DB 12; Length 297;
Best Local Similarity 89.3%; Pred. No. 5.8e-93;
Matches 267; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANKHVLTTLAIQDATSQIKNTTPTYLTPQDGLGSPSNLSEITTSQTTILASTTP 120
DB 61 FIASANKHVLTTLAIQDATSQIKNTTPTYLTPQDGLGSPSNLSEITTSQTTILASTTP 120

QY 121 GVKSNLQPTTVKTKNTTTTQTPSKPTTKQKQKPPKPNNDFFHFEVFNVPVCSICSNP 180
DB 121 GVKSTLQSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNNDFFHFEVFNVPVCSICSNP 180

QY 181 TCWAIKRIKPKKPGKKTITTKPTKPKTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240
DB 181 TCWAIKRIKPKKPGKKTITTKPTKPKTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240

Query Match      88.3%; Score 1364; DB 12; Length 297;
Best Local Similarity 89.3%; Pred. No. 5.8e-93;
Matches 267; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANKHVLTTLAIQDATSQIKNTTPTYLTPQDGLGSPSNLSEITTSQTTILASTTP 120
DB 61 FIASANKHVLTTLAIQDATSQIKNTTPTYLTPQDGLGSPSNLSEITTSQTTILASTTP 120

QY 121 GVKSNLQPTTVKTKNTTTTQTPSKPTTKQKQKPPKPNNDFFHFEVFNVPVCSICSNP 180
DB 121 GVKSTLQSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNNDFFHFEVFNVPVCSICSNP 180

QY 181 TCWAIKRIKPKKPGKKTITTKPTKPKTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240
DB 181 TCWAIKRIKPKKPGKKTITTKPTKPKTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240

```

```

Db      181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKSKKEVPTTKPTBPTINTTK 240
QY      241 TMTITLLTNNTGNPKLSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 297
Db      241 TMTITLLTSTNTRNPELTSTOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTSR 297

RESULT 7
Q9YVB4 PRELIMINARY; PRT; 298 AA.
AC Q9YVB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Attachment glycoprotein G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WV5222;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932(1998).
DR EMBL; AF065406; AAD02942.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
SQ SEQUENCE 298 AA; 32781 MW; AE7C9E05547C5745 CRC64;

Query Match 87.8%; Score 1356; DB 12; Length 298;
Best Local Similarity 88.9%; Pred. No. 2.3e-92;
Matches 265; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLITAI 60
Db 1 MSKNKQORTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLITAI 60

QY 61 FIASANKVTLTALIODATSOIKNTPTTYPOLGIGSPNLSHITSQTITLASTTP 120
Db 61 FIASANKVTLTALIODATSOIKNTPTTYPOLGIGSPNLSHITSQTITLASTTP 120

QY 121 GVKSNLOPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFHFEVNFVPCISCSNNP 180
Db 121 SVESTLQSTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFHFEVNFVPCISCSNNP 180

QY 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKSKKEVPTTKPTBPTINTTK 240
Db 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKSKKEVPTTKPTBPTINTTK 240

QY 241 TMTITLLTNNTGNPKLSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 298
Db 241 PNIRITLLTNTTSGPSTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 298

RESULT 8
Q82071 PRELIMINARY; PRT; 297 AA.
AC Q82071;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbiza J., Fabrasile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459(1994).
DR EMBL; Z33432; CAAB3875.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
SQ SEQUENCE 297 AA; 32555 MW; 9D0A69BECAE8B6C8 CRC64;

Query Match 87.7%; Score 1354; DB 12; Length 297;
Best Local Similarity 89.2%; Pred. No. 3.2e-92;
Matches 265; Conservative 10; Mismatches 22; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLITAI 60
Db 1 MSKNKQORTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLITAI 60

QY 61 FIASANKVTLTALIODATSOIKNTPTTYPOLGIGSPNLSHITSQTITLASTTP 120
Db 61 FIASANKVTLTALIODATSOIKNTPTTYPOLGIGSPNLSHITSQTITLASTTP 120

QY 121 GVKSNLOPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFHFEVNFVPCISCSNNP 180
Db 121 GVKSNLOPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFHFEVNFVPCISCSNNP 180

QY 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKSKKEVPTTKPTBPTINTTK 240
Db 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKSKKEVPTTKPTBPTINTTK 240

QY 241 TMTITLLTNNTGNPKLSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 297
Db 241 TMTITLLTNTTRNPELTSTOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTSR 297

RESULT 9
Q82067 PRELIMINARY; PRT; 297 AA.
AC Q82067;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbiza J., Fabrasile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459(1994).
DR EMBL; Z33432; CAAB3875.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
SQ SEQUENCE 297 AA; 32555 MW; 9D0A69BECAE8B6C8 CRC64;

```

Query Match 87.1%; Score 1345; DB 12; Length 297;  
Best Local Similarity 88.9%; Pred. No. 1.5e-91;  
Matches 264; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIIITAI 60  
DB 1 MSKNKQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIIIAII 60

QY 61 FIASANHKVLTITAIIOQATSOIKNTTPTLTQDPQLGISFNSLSEITSOITLILASTTP 120  
DB 61 FIASANHKVLTITAIIOQATSOIKNTTPTLTQDPQLGISFNSLSEITSLITLIDSTTP 120

QY 121 GVKSNLQPTTVKTKNTTPTQPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180  
DB 121 GVKSTLQSTTVGKNTTPTQPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180

QY 181 TCWAICKRIPNKKPGKKTITTKPTKPTKTKKOLKPKQTTKPKVEPTTKPTTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKKTITTKPTKPTKTKKOLKPKQTTKPKVEPTTKPTTEPTINTTK 240

QY 241 TNITITLLTNNTGNPKLTSMQETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 297  
DB 241 TNITITLLTNNTGNPKLTSMQETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 297

RESULT 10  
Q82058 PRELIMINARY; PRT; 297 AA.

AC Q82058; (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE (MAD-1-89) subgroup A, G glycoprotein.  
GN G.

OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=subgroup A;  
RX MEDLINE=94335057; PubMed=8057427;  
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabraile S., Russi J.,  
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
RA Melero J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): cocirculating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein.";  
RL J. Virol. 68:5448-5459(1994).  
DR EMBL: Z33456; CAA83879.1; --  
DR InterPro; IPR000925; Glycoprot.G.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.  
SQ SEQUENCE 297 AA; 32741 MW; 0B567A174BF64964 CRC64;

Query Match 87.0%; Score 1344; DB 12; Length 297;  
Best Local Similarity 88.5%; Pred. No. 1.7e-91;  
Matches 262; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIIITAI 60  
DB 1 MSKNKQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIIIAII 60

QY 61 FIASANHKVLTITAIIOQATSOIKNTTPTLTQDPQLGISFNSLSEITSOITLILASTTP 120  
DB 61 FIASANHKVLTITAIIOQATSOIKNTTPTLTQDPQLGISFNSLSEITSOITLILASTTP 120

QY 121 GVKSNLQPTTVKTKNTTPTQPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180  
DB 121 SAESTPQSTTVKTKNTTPTQPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180

QY 181 TCWAICKRIPNKKPGKKTITTKPTKPTKTKKOLKPKQTTKPKVEPTTKPTTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKKTITTKPTKPTKTKKOLKPKQTTKPKVEPTTKPTTEPTINTTK 240

QY 241 TNITITLLTNNTGNPKLTSMQETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 296  
DB 241 TNITITLLTNNTGNPKLTSMQETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 296

RESULT 11  
Q91947 PRELIMINARY; PRT; 297 AA.

AC Q91947; (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Attachment protein.  
GN G.

OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=CH34;  
RX MEDLINE=38418504; PubMed=9747732;  
RA Peret T.C.F., Hall C.B., Schnabel K.C., Golub J.A., Anderson L.J.;  
RT "Circulation Patterns of Genetically Distinct Group A and B Strains of  
RT Human Respiratory Syncytial Virus in a Community.";  
RL J. Gen. Virol. 79:2221-2229(1998).  
DR EMBL: AF065257; AAC36327.1; --  
DR InterPro; IPR000925; Glycoprot.G.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.  
SQ SEQUENCE 297 AA; 32564 MW; 377C7CBCDD0A3645 CRC64;

Query Match 86.7%; Score 1339; DB 12; Length 297;  
Best Local Similarity 88.2%; Pred. No. 4.1e-91;  
Matches 262; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIIITAI 60  
DB 1 MSKNKQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIIIAII 60

QY 61 FIASANHKVLTITAIIOQATSOIKNTTPTLTQDPQLGISFNSLSEITSOITLILASTTP 120  
DB 61 FIASANHKVLTITAIIOQATSOIKNTTPTLTQDPQLGISFNSLSEITSLITLIDSTTP 120

QY 121 GVKSNLQPTTVKTKNTTPTQPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180  
DB 121 GVKSTLQSTTVGKNTTPTQPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180

QY 181 TCWAICKRIPNKKPGKKTITTKPTKPTKTKKOLKPKQTTKPKVEPTTKPTTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKKTITTKPTKPTKTKKOLKPKQTTKPKVEPTTKPTTEPTINTTK 240

QY 241 TNITITLLTNNTGNPKLTSMQETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 297  
DB 241 TDITTFPSNTTNPENLTSQMETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 297

RESULT 12  
Q9YVB3 PRELIMINARY; PRT; 293 AA.

AC Q9YVB3; (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Attachment glycoprotein G (fragment).  
GN Human respiratory syncytial virus.  
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

```

OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WV6973;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932 (1998).
DR EMBL; AF065407; AAD02943.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
FT NON TER 1
SQ SEQUENCE 293 AA; 32125 MW; 17BSB43396A63CCF CRC64;

Query Match 86.4%; Score 1336; DB 12; Length 293;
Best Local Similarity 90.1%; Pred. No. 6.4e-91;
Matches 264; Conservative 7; Mismatches 22; Indels 0; Gaps 0;

QY 6 DQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPIA 65
DB 1 DQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPIA 60

QY 66 NHKVTLTALIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTPGVKN 125
DB 61 NHKVTSTTTIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTPGVKST 120

QY 126 LQPTTKTKNTTTTQTPSKPTTKQKPNKPNNDPHEVFVPCISCSNNPTCWA 195
DB 121 LQSTTVTKNTTTTQTPSKPTTKQKPNKPNNDPHEVFVPCISCSNNPTCWA 180

QY 186 CKRIPNKKPKGKTKTKPTTKTKKDKLPQTKPKKEVPTTKPEEPTINTKNIT 245
DB 181 CKRIPNKKPKGKTKTKPTTKTKKDKLPQTKPKKEVPTTKPEEPTINTKNIT 240

QY 246 TLTNTNTGNPKLTSMQETFSSTSEGNLSPSQVSTTSEHPSPSPSPNTTRQ 298
DB 241 TLTNTNTRNPELTSMQETFSSTSEGNLSPSQVSTTSEHPSPSPSPNTTRQ 293

RESULT 13
Q86359
ID Q86359 PRELIMINARY; PRT; 295 AA.
AC Q86359;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE G protein (Fragment).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RA Cane P.A.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RX EMBL; X73353; CAA51764.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
FT NON TER 1
SQ SEQUENCE 295 AA; 32306 MW; 671DCF044157D284 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 295;
Best Local Similarity 88.5%; Pred. No. 9.4e-91;
Matches 261; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 3 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 62
DB 1 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 60

QY 63 ASANHKVTLTALIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTPGV 122
DB 61 ASANHKITSTTTIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTPGV 120

```

```

FT NON TER 1
SQ SEQUENCE 295 AA; 32490 MW; 48D835F670FF8006 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 295;
Best Local Similarity 88.4%; Pred. No. 9.4e-91;
Matches 260; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY 3 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 62
DB 1 KTKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 60

QY 63 ASANHKVTLTALIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTPGV 122
DB 61 ASANHKVTLTALIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTPGV 120

QY 123 KSNLQPTTKTKNTTTTQTPSKPTTKQKPNKPNNDPHEVFVPCISCSNNPTC 182
DB 121 ESTPOSTTVTKNTTTTQTPSKPTTKQKPNKPNNDPHEVFVPCISCSNNPTC 180

QY 183 WAICKRIPNKKPKGKTKTKPTTKTKKDKLPQTKPKKEVPTTKPEEPTINTKN 242
DB 181 WAICKRIPNKKPKGKTKTKPTTKTKKDKLPQTKPKKEVPTTKPEEPTINTKN 240

QY 243 ITTLLTNTNTPKLTSMQETFSSTSEGNLSPSQVSTTSEHPSPSPSPNTT 296
DB 241 ITTLLTNTNTPKLTSMQETFSSTSEGNLSPSQVSTTSEHPSPSPSPNTT 294

RESULT 14
Q86360
ID Q86360 PRELIMINARY; PRT; 295 AA.
AC Q86360;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE G protein (Fragment).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-642;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-642;
RA Cane P.A.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RX EMBL; X73354; CAA51765.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
FT NON TER 1
SQ SEQUENCE 295 AA; 32306 MW; 671DCF044157D284 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 295;
Best Local Similarity 88.5%; Pred. No. 9.4e-91;
Matches 261; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 3 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 62
DB 1 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 60

QY 63 ASANHKVTLTALIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTPGV 122
DB 61 ASANHKITSTTTIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTPGV 120

```

```

QY 123 KSNLOPTVTKNTTITOTOPSKPTTKQKONKPPKNPNNDFFHEVFNFPVPCSCSNPTC 182
DB 121 KSTLQSTTVGTNTTTOAQPNKPTTKQKONKPPKNPNNDFFHEVFNFPVPCSCSNPTC 180
QY 183 WAICKRIPNKKPGKTKTKPTTKKPTTKKOLKPKQTKPKKVPPTTKPTTEPTINTTKN 242
DB 181 WAICKRIPNKKPGKTKTKPTTKKPTTKKOLKPKQTKPKKVPPTTKPTTEPTINTTKN 240
QY 243 ITTTLTNNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPSPPTTR 297
DB 241 ITTTLTNNTRNPELTSMQETHSTSSGNLSPSQVSTTSBHPSPSPSPPTTR 295

RESULT 15
ID Q82065 PRELIMINARY; PRT; 298 AA.
AC Q82065;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein.
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrasile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5455(1994).
DR EMBL; Z33426; CAA83869.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PROSHPANTHEINE; 1.
SQ SEQUENCE 298 AA; 32729 MW; 4E890FDD43B45744 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 298;
Best Local Similarity 88.2%; Pred. No. 9.5e-91;
Matches 261; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKTDNLNHLFISSGLYKLNKLSVAQITLSILAMISTSLIITAI 60
DB 1 MSKTKDQRTAKTLEKTDNLNHLFISSGLYKLNKLSVAQITLSILAMISTSLIITAI 60
QY 61 FIASANKHVKLTTLTIIQDATSQIKNTPTLTODPOLGISFNSLSEITSOITTLASTTP 120
DB 61 FIASANKHVKLTTLTIIQDATSQIKNTPTLTQNPQLGISFNSLSEITSOITTLASTTP 120
QY 121 GVKSNLOPTVTKNTTITOTOPSKPTTKQKONKPPKNPNNDFFHEVFNFPVPCSCSNNP 180
DB 121 SAESTPLSTTVTKNTTITOTOPSKPTTKQKONKPPKNPNNDFFHEVFNFPVPCSCSNNP 180
QY 181 TCWAICKRIPNKKPGKTKTKPTTKKPTTKKOLKPKQTKPKKVPPTTKPTTEPTINTTK 240
DB 181 TCWAICKRIPNKKPGKTKTKPTTKKPTTKKOLKPKQTKPKKVPPTTKPTTEPTINTTK 240
QY 241 TNIITTLTNNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPSPPTTR 296
DB 241 TNIITTLTNNTRNPELTSMQETHSTSSGNLSPSQVSTTSBHPSPSPSPPTTR 296

```

Search completed: October 29, 2003, 17:40:52  
Job time : 75.283 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:12:25 ; Search time 16.3057 Seconds  
(without alignments)  
859.454 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKQRTAKLEKTWDL.....VSTSEHPSPSPPTTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 1530  | 99.1        | 298    | 1     | VGLG_HRSVL  |
| 2          | 1418  | 91.8        | 298    | 1     | VGLG_HRSVA  |
| 3          | 1343  | 87.0        | 297    | 1     | VGLG_HRSV3  |
| 4          | 1343  | 87.0        | 297    | 1     | VGLG_HRSV6  |
| 5          | 1328  | 86.0        | 297    | 1     | VGLG_HRSV7  |
| 6          | 1324  | 85.8        | 297    | 1     | VGLG_HRSV4  |
| 7          | 1313  | 85.0        | 298    | 1     | VGLG_HRSV5  |
| 8          | 1307  | 84.7        | 297    | 1     | VGLG_HRSV2  |
| 9          | 767.5 | 49.7        | 292    | 1     | VGLG_HRSV8  |
| 10         | 761   | 49.3        | 292    | 1     | VGLG_HRSV1  |
| 11         | 309.5 | 20.0        | 263    | 1     | VGLG_HRSV1  |
| 12         | 308.5 | 20.0        | 263    | 1     | VGLG_HRSV4  |
| 13         | 299   | 19.4        | 257    | 1     | VGLG_HRSV2  |
| 14         | 295.5 | 19.1        | 263    | 1     | VGLG_HRSV2  |
| 15         | 294.5 | 19.1        | 263    | 1     | VGLG_HRSV2  |
| 16         | 285   | 18.5        | 263    | 1     | VGLG_HRSV5  |
| 17         | 284   | 18.4        | 263    | 1     | VGLG_HRSV7  |
| 18         | 279   | 18.1        | 257    | 1     | VGLG_HRSV7  |
| 19         | 270   | 17.5        | 257    | 1     | VGLG_HRSVL  |
| 20         | 267   | 17.3        | 257    | 1     | VGLG_HRSV7  |
| 21         | 208   | 13.5        | 307    | 1     | SGS3_DROME  |
| 22         | 198.5 | 12.9        | 2812   | 1     | ZAN_HUMAN   |
| 23         | 194   | 12.6        | 662    | 1     | MUC1_XENLA  |
| 24         | 193   | 12.5        | 5179   | 1     | MUC2_HUMAN  |
| 25         | 191   | 12.4        | 1161   | 1     | DAN4_YEAST  |
| 26         | 183   | 11.9        | 217    | 1     | SGS3_DROS1  |
| 27         | 171.5 | 11.1        | 2282   | 1     | ZAN_RABIT   |
| 28         | 170   | 11.0        | 5703   | 1     | MUSE_HUMAN  |
| 29         | 169.5 | 11.0        | 1357   | 1     | AMTH_YEAST  |
| 30         | 164.5 | 10.7        | 2476   | 1     | ZAN_PIG     |
| 31         | 160   | 10.4        | 263    | 1     | SGS3_DROYA  |
| 32         | 159.5 | 10.3        | 907    | 1     | VGP3_EBV    |
| 33         | 157.5 | 10.2        | 5376   | 1     | ZAN_MOUSE   |

ALIGNMENTS

RESULT 1

| ID | VGLG_HRSVL   | STANDARD; | PRT; | 298 AA. |
|----|--|-----------|------|---------|
| AC | P20895;  |           |      |         |
| DT | 01-FEB-1991 (Rel. 17, Created)   |           |      |         |
| DT | 01-FEB-1996 (Rel. 33, Last sequence update)  |           |      |         |
| DT | 01-FEB-1996 (Rel. 33, Last annotation update)  |           |      |         |
| DE | Major surface glycoprotein G (Attachment glycoprotein G).  |           |      |         |
| GN | G.   |           |      |         |
| OS | Human respiratory syncytial virus (subgroup A / strain Long).  |           |      |         |
| OC | Viruses; ssRNA negative-strand viruses; Mononegavirales;   |           |      |         |
| OC | Paramyxoviridae; Pneumovirinae; Pneumovirus.   |           |      |         |
| OX | NCBI_TaxID=11260;  |           |      |         |
| RN | [1]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=87289657; PubMed=2441388;  |           |      |         |
| RA | Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;  |           |      |         |
| RT | "The G glycoprotein of human respiratory syncytial viruses of  |           |      |         |
| RT | subgroups A and B: extensive sequence divergence between   |           |      |         |
| RT | antigenically related proteins.";  |           |      |         |
| RL | Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).  |           |      |         |
| CC | -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  |           |      |         |
| CC | RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND   |           |      |         |
| CC | HEMAGGLUTININATING ACTIVITIES.   |           |      |         |
| CC | -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED   |           |      |         |
| CC | CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.   |           |      |         |
| CC | -!- FTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS   |           |      |         |
| CC | DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  |           |      |         |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration   |           |      |         |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation -  |           |      |         |
| CC | the European Bioinformatics Institute. There are no restrictions on its  |           |      |         |
| CC | use by non-profit institutions as long as its content is in no way   |           |      |         |
| CC | modified and this statement is not removed. Usage by and for commercial  |           |      |         |
| CC | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |           |      |         |
| CC | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).  |           |      |         |
| DR | EMBL; M17212; AAA47411.1; --   |           |      |         |
| DR | PIR; A32703; MGNZRL.   |           |      |         |
| DR | InterPro; IPR000925; Glycoprot G.  |           |      |         |
| DR | Pfam; PF00802; Glycoprotein G; 1.  |           |      |         |
| KW | Transmembrane; Glycoprotein.   |           |      |         |
| FT | DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).   |           |      |         |
| FT | TRANSMEM 38 66 POTENTIAL.  |           |      |         |
| FT | DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).   |           |      |         |
| FT | CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).   |           |      |         |
| FT | CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).   |           |      |         |
| FT | CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).   |           |      |         |
| FT | CARBOHYD 250 250 N-LINKED (GLCNAC. .) (POTENTIAL).   |           |      |         |
| FT | CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).   |           |      |         |
| FT | CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).   |           |      |         |
| SQ | SEQUENCE 298 AA; 32781 MW; B79FEFA4A473B0E CRC64;  |           |      |         |

Query Match 99.1%; Score 1530; DB 1; Length 298;  
Best Local Similarity 99.0%; Pred. No. 8.6e-93;  
Matches 295; Conservative 1; Mismatches 2; Indels 0; Gaps 0;





```

CC CC HEMAGGLUTININATING ACTIVITIES.
CC CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC CC PIR; JQ1205; JQ1205.
CC CC DR InterPro; IPR000925; Glycoprot G.
CC CC DR Pfam; PF00802; Glycoprotein G; 1.
CC CC KW Transmembrane; Glycoprotein.
CC CC FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
CC CC FT TRANSMEM 38 66 POTENTIAL.
CC CC FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
CC CC FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC SQ SEQUENCE 297 AA; 32525 MW; 48448F9E091E1802 CRC64;

Query Match 87.0%; Score 1343; DB 1; Length 297;
Best Local Similarity 88.6%; Pred. No. 1.2e-80;
Matches 263; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKNKDORTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNKDORTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIIVAI 60
QY 61 FIASANKHVTLLTAIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTTTILASTTP 120
DB 61 FIASANKHVTLLTAIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTTTILASTTP 120
QY 121 GVKSNLPQTTVKTKNTTTTQPSKPTTKQRQKPNKPNNDHFEVFNVPSCISNNP 180
DB 121 GVKSTLQSTTVGTVKNTTTTQAPNKPPTTKQRQKPNKPNNDHFEVFNVPSCISNNP 180
QY 181 TCWAICKRIENKPKGKTTTKPKPTKTKDLPQTTKPKKEVPTTKPTBPTINTTK 240
DB 181 TCWAICKRIENKPKGKTTTKPKPTKTKDLPQTTKPKKEVPTTKPTBPTINTTK 240
QY 241 TNITITLLTNNTGNPKLTSQMETFSTSGEGLSPSQSVSTTSEHPSQSPSPNTTR 297
DB 241 TNITITLLTNNTGNPKLTSQMETFSTSGEGLSPSQSVSTTSEHPSQSPSPNTTR 297

RESULT 4
VGLG_HRSV6 STANDARD; PRT; 297 AA.
AC P27025;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb6256).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11256;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC CC PIR; JQ1208; JQ1208.
CC CC DR InterPro; IPR000925; Glycoprot G.
CC CC DR Pfam; PF00802; Glycoprotein G; 1.
CC CC KW Transmembrane; Glycoprotein.

```

```

FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA; 32708 MW; 6781756C38B64A80 CRC64;

Query Match 87.0%; Score 1343; DB 1; Length 297;
Best Local Similarity 88.5%; Pred. No. 1.2e-80;
Matches 262; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY 1 MSKNKDORTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNKDORTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
QY 61 FIASANKHVTLLTAIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTTTILASTTP 120
DB 61 FIASANKHVTLLTAIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTTTILASTTP 120
QY 121 GVKSNLPQTTVKTKNTTTTQPSKPTTKQRQKPNKPNNDHFEVFNVPSCISNNP 180
DB 121 SAESTPQSTTVKTKNTTTTQIQPSKPTTKQRQKPNKPNNDHFEVFNVPSCISNNP 180
QY 181 TCWAICKRIENKPKGKTTTKPKPTKTKDLPQTTKPKKEVPTTKPTBPTINTTK 240
DB 181 TCWAICKRIENKPKGKTTTKPKPTKTKDLPQTTKPKKEVPTTKPTBPTINTTK 240
QY 241 TNITITLLTNNTGNPKLTSQMETFSTSGEGLSPSQSVSTTSEHPSQSPSPNTT 296
DB 241 TNITITLLTNNTGNPEYTSQETLSTSPGPNPSPQSVTTSEYSPSPSPNTT 296

RESULT 5
VGLG_HRSV7 STANDARD; PRT; 297 AA.
AC P27036;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb6614).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11257;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC CC PIR; JQ1209; JQ1209.
CC CC DR InterPro; IPR000925; Glycoprot G.
CC CC DR Pfam; PF00802; Glycoprotein G; 1.
CC CC KW Transmembrane; Glycoprotein.
CC CC FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
CC CC FT TRANSMEM 38 66 POTENTIAL.
CC CC FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
CC CC FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).

```



```

Db 61 FIASANKVLTATTAIQDATSOIKNTPTTYLTQNPOLGISFFNLSGTTSQTALITTP 120
QY 121 GVKSNLQPTTKVTKNTTTTQTOPSKPTTKORQKNPKPNNDHFEVFNVPVCSICSNPP 180
Db 121 SVESILQSTTVKNTNTTQIQSKPTTKORQKNPKPNNDHFEVFNVPVCSICSNPP 180
QY 181 TCWAIKCRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
Db 181 TCWAIKCRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
QY 241 TNITTTLLTNNTTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHSPSQSSPNTTTR 298
Db 241 PNIRTTLLTNSTTGNLEHTSQEBTSLHSTSEGNLSPSQVSTTSEHSPSQSSPNTTTR 298

RESULT 8
VGLG_HRSV2
ID VGLG_HRSV2 STANDARD; PRT; 297 AA.
AC P27021;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb642).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11252;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- FTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR; JQ1204; JQ1204.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA; 3745 MW; FC72A7F3A8BFF67C CRC64;

Query Match 84.7%; Score 1307; DB 1; Length 297;
Best Local Similarity 86.5%; Pred. No. 2.6e-78;
Matches 257; Conservative 12; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKTDWTLNHLFFISSGLYKLNLSVAQITLSILAMISLITAI 60
Db 1 MSKNKQORTAKTLEKTDWTLNHLFFISSGLYKLNLSVAQITLSILAMISLITAI 60
QY 61 FIASANKVLTATTAIQDATSOIKNTPTTYLTQNPOLGISFFNLSGTTSQTALITTP 120
Db 61 FIASANKVLTATTAIQDATSOIKNTPTTYLTQNPOLGISFFNLSGTTSQTALITTP 120
QY 121 GVKSNLQPTTKVTKNTTTTQTOPSKPTTKORQKNPKPNNDHFEVFNVPVCSICSNPP 180
Db 121 GVKLTQSTTVKNTNTTQIQSKPTTKORQKNPKPNNDHFEVFNVPVCSICSNPP 180
QY 181 TCWAIKCRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
Db 181 TCWAIKCRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240

```

```

QY 241 TNITTTLLTNNTTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHSPSQSSPNTTTR 297
Db 241 TNITTTLSNTARNPELTISQMETFHSTSEGNLSPSQVSTTSEHSPSQSSPNTTTR 297

RESULT 9
VGLG_HRSV8
ID VGLG_HRSV8 STANDARD; PRT; 292 AA.
AC P23041;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup B / strain 8/60).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11258;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90357765; PubMed=1697126;
RA Sullender W.M., Anderson K., Wertz G.W.;
RT "The respiratory syncytial virus subgroup B attachment glycoprotein:
RT analysis of sequence, expression from a recombinant vector, and
RT evaluation as an immunogen against homologous and heterologous
RT subgroup virus challenge.";
RL Virology 178:195-203 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374595; PubMed=1895391;
RA Sullender W.M., Mufson M.M., Anderson L.J., Wertz G.W.;
RT "Genetic diversity of the attachment protein of subgroup B
RT respiratory syncytial viruses.";
RL J. Virol. 65:5425-5434 (1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- FTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; M55633; AAA47413.1; -.
DR EMBL; M73545; AAA47408.1; -.
DR PIR; A37077; MGNZ60.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 292 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 292 AA; 32143 MW; 8EC60C85EF057BB5 CRC64;

Query Match 49.7%; Score 767.5; DB 1; Length 292;
Best Local Similarity 55.3%; Pred. No. 2.7e-43;
Matches 162; Conservative 34; Mismatches 94; Indels 3; Gaps 3;

QY 1 MSKNKQORTAKTLEKTDWTLNHLFFISSGLYKLNLSVAQITLSILAMISLITAI 60
Db 1 MSKNKQORTASTLEKTDWTLNHLFIVISSCLYKLNLSVAQITLSILAMISLITAI 60

```



```

Query Match      20.0%; Score 308.5; DB 1; Length 263;
Best Local Similarity 32.5%; Pred. No. 1.5e-13;
Matches 90; Conservative 32; Mismatches 116; Indels 39; Gaps 7;

QY 1 MSKNKQORTAKTLEKWTDTLHLLFISGLYKMLKSAQITLSILAMIISTSLIITAI 60
DB 1 MSNTHHPKFKTLKRAWKASKYFIVGLSKLYKFNLSIVQTLTTLAMITLSLITAI 60
QY 61 FIASANKHVTLTALIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTP 120
DB 1 MSNTHHPKFKTLKRAWKASKYFIVGLSKLYKFNLSIVQTLTTLAMITLSLITAI 60
QY 61 FIASANKHVTLTALIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTP 120
DB 1 YISVGNKAKPTSKPTTQOQLOQNHPTPLTEHNY-----KSTHTSIQSTTL 108
QY 121 GVKNSLOPTTKVTKNTTQTQPSK---PTTKQKQKPPNKP-----NNDHF--FEVP 168
DB 109 SQPNIDTTSSTGYGHTNTRQNKIKSQSTPLATRAPPINPLGSPNPNHQNNSQTL 168
QY 169 NFVPCSCISNPTCWAIC-----KRIPN-----KPKGKTKTKTKTKTKTKKDLK 216
DB 169 PHVPCSTCEGNPACSPLCQIELERAPSSAPTITLKKAPKTKTKTKTKTKTKTSPEAK 228
QY 217 POTTKEVPTTKPTTEPTINTTKNTITLLTNNTT 253
DB 229 LQTKKIMATPQOGILSSP---EHQTNQSTTQISQHTS 262

RESULT 13
VGLG BRSV4
ID VGLG BRSV4 STANDARD; PRT; 263 AA.
AC Q10684;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Bovine respiratory syncytial virus (strain 4642) (BRS).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=82820;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97288324; PubMed=9143302;
RA Furze J., Roberts S., Wertz G., Taylor G.;
RT "Antigenically distinct G glycoproteins of BRSV strains share a high
RL degree of genetic homogeneity.";
RL Virology 231:48-58(1997).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- SIMILARITY: TO THE G PROTEINS OF EITHER THE SUBGROUP A OR B
CC HRS VIRUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y08718; CAAG9968.1; --
CC HSSP; P22261; 1BRV.
CC InterPro; IPR000925; Glycoprot G.
CC Pfam; PF00802; Glycoprotein G; 1.
CC Transmembrane; Glycoprotein.
CC DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 38 66 POTENTIAL.
CC DOMAIN 67 263 EXTRACELLULAR (POTENTIAL).
CC DISULFID 173 186 BY SIMILARITY.
CC FT DISULFID 176 182 BY SIMILARITY.
CC FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 163 163 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 251 251 N-LINKED (GLCNAC...) (POTENTIAL).
CC SEQUENCE 263 AA; 28976 MW; D592D79EF8EE9535 CRC64;

QY 1 MSKNKQORTAKTLEKWTDTLHLLFISGLYKMLKSAQITLSILAMIISTSLIITAI 60
DB 1 MSNTHHPKFKTLKRAWKASKYFIVGLSKLYKFNLSIVQTLTTLAMITLSLITAI 60
QY 61 FIASANKHVTLTALIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTP 120
DB 1 MSNTHHPKFKTLKRAWKASKYFIVGLSKLYKFNLSIVQTLTTLAMITLSLITAI 60
QY 61 FIASANKHVTLTALIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTP 120
DB 1 YISVGNKAKPTSKPTTQOQLOQNHPTPLTEHNY-----KSTHTSIQSTTL 108
QY 121 GVKNSLOPTTKVTKNTTQTQPSK---PTTKQKQKPPNKP-----NNDHF--FEVP 168
DB 109 SQPNIDTTSSTGYGHTNTRQNKIKSQSTPLATRAPPINPLGSPNPNHQNNSQTL 168
QY 169 NFVPCSCISNPTCWAIC-----KRIPN-----KPKGKTKTKTKTKTKTKKDLK 216
DB 169 PHVPCSTCEGNPACSPLCQIELERAPSSAPTITLKKAPKTKTKTKTKTKTKTSPEAK 228
QY 217 POTTKEVPTTKPTTEPTINTTKNTITLLTNNTT 253
DB 229 LQTKKIMATPQOGILSSP---EHQTNQSTTQISQHTS 262

RESULT 13
VGLG BRSV4
ID VGLG BRSV4 STANDARD; PRT; 257 AA.
AC P22261;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Bovine respiratory syncytial virus (strain Copenhagen) (BRS).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11248;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91012801; PubMed=2214024;
RA Lerch R.A., Anderson K., Wertz G.W.;
RT "Nucleotide sequence analysis and expression from recombinant vectors
RT demonstrate that the attachment protein G of bovine respiratory
RT syncytial virus is distinct from that of human respiratory syncytial
RT virus.";
RL J. Virol. 64:5559-5569(1990).
RN [2]
RP STRUCTURE BY NMR OF 158-189.
RC STRAIN=391-2;
RX MEDLINE=97098087; PubMed=8942628;
RA Doreleijers J.F., Langedijk J.P.M., Haard K., Boelens R.,
RA Rullmann J.A., Schaaper W.M., van Oirschot J.T., Kaptein R.;
RT "Solution structure of the immunodominant region of protein G of
RT bovine respiratory syncytial virus.";
RL Biochemistry 35:14684-14688(1996).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- SIMILARITY: TO THE G PROTEINS OF EITHER THE SUBGROUP A OR B
CC HRS VIRUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M58307; AAA42810.1; --

```







Search completed: October 29, 2003, 17:38:37  
Job time : 18.3057 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:33:35 ; Search time 28.6755 Seconds  
(without alignments)  
999.400 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKQRTAKTLEKWTDL.....VSTTSHPSQSPSPNTTQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 1530  | 99.1        | 298    | 1 MGNZRL | major surface glyc |
| 2          | 1418  | 91.8        | 298    | 1 MGNZ   | major surface glyc |
| 3          | 1343  | 87.0        | 297    | 2 JQ1205 | attachment protein |
| 4          | 1343  | 87.0        | 297    | 2 JQ1208 | attachment protein |
| 5          | 1333  | 86.3        | 298    | 2 JQ5680 | G protein - Human  |
| 6          | 1328  | 86.0        | 297    | 2 JQ1209 | attachment protein |
| 7          | 1324  | 85.8        | 297    | 2 JQ1206 | attachment protein |
| 8          | 1313  | 85.0        | 298    | 2 JQ1207 | attachment protein |
| 9          | 1307  | 84.7        | 297    | 2 JQ1204 | attachment protein |
| 10         | 767.5 | 49.7        | 292    | 1 MGNZ60 | major surface glyc |
| 11         | 761   | 49.3        | 292    | 1 MGNZ18 | major surface glyc |
| 12         | 425   | 27.5        | 467    | 1 VHNZ   | nucleosid prote    |
| 13         | 299   | 19.4        | 250    | 2 P00768 | glycoprotein G - b |
| 14         | 299   | 19.4        | 257    | 1 MGNZR  | major surface glyc |
| 15         | 295.5 | 19.1        | 263    | 2 JQ2284 | glycoprotein G - b |
| 16         | 284   | 18.4        | 263    | 2 JQ2388 | glycoprotein G - o |
| 17         | 280   | 18.1        | 248    | 2 P00769 | glycoprotein G - b |
| 18         | 272   | 17.6        | 263    | 2 A48732 | attachment glycop  |
| 19         | 208   | 13.5        | 307    | 1 GSPF3  | salivary glue prot |
| 20         | 194   | 12.6        | 662    | 2 A45155 | mucin FIM-C.1 - Af |
| 21         | 193   | 12.5        | 3020   | 2 A43932 | mucin 2 precursor, |
| 22         | 191   | 12.4        | 851    | 2 T22696 | hypothetical prote |
| 23         | 191   | 12.4        | 1161   | 2 S57180 | probable membrane  |
| 24         | 190   | 12.3        | 322    | 2 A53715 | apomucin precursor |
| 25         | 188   | 12.2        | 1832   | 2 T31113 | mucin-like glycop  |
| 26         | 185.5 | 12.0        | 379    | 2 S50125 | larval glue protei |
| 27         | 183   | 11.9        | 217    | 2 S01358 | salivary glue prot |
| 28         | 183   | 11.9        | 354    | 2 T46740 | microfilarial shea |
| 29         | 177   | 11.5        | 770    | 2 T22808 | hypothetical prote |

30 175 11.3 825 2 T29634  
31 172.5 11.2 327 2 S20074  
32 172.5 11.2 1777 2 T34369  
33 171.5 11.1 216 2 151920  
34 170.5 11.0 279 2 S53363  
35 170 11.0 3570 2 T45025  
36 169.5 11.0 1367 1 S48478  
37 168 10.9 402 2 E86185  
38 167.5 10.8 294 2 A37232  
39 167.5 10.8 1630 2 A53577  
40 164.5 10.7 2476 2 T34022  
41 163 10.6 371 2 S20075  
42 163 10.6 477 2 S53362  
43 160.5 10.4 393 2 S62335  
44 160.5 10.4 866 2 T45462  
45 160.5 10.4 867 2 T45463

#### ALIGNMENTS

##### RESULT 1

MGNZRL

major surface glycoprotein G - human respiratory syncytial virus (strain Long)

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999

C:Accession: A32703; S12279

R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.

Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987

A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and

A:Reference number: A32703; PMID:87289657; PMID:2441388

A:Accession: A32703

A:Molecule type: mRNA

A:Residues: 1-298 <OH>

A:Cross-references: GB:M17212; NID:G333940; PIDN:AAA47411.1; PID:G333941

R:Garcia-Barreno, B.; Portela, A.; Delgado, T.; Lopez, J.A.; Melero, J.A.

EMBO J. 9, 4181-4187, 1990

A:Title: Frame shift mutations as a novel mechanism for the generation of neutralizati

A:Reference number: S12279; PMID:91065351; PMID:2249671

A:Accession: S12279

A:Molecule type: mRNA

A:Residues: 1-298 <GAR>

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:41-63/Domain: transmembrane #status predicted <TMN>

F:85,103,135,179,237,250,251,273,294/Binding site: carbohydrate (Asn) (covalent) #stat

Query Match 99.1%; Score 1530; DB 1; Length 298;

Best Local Similarity 99.0%; Pred. No. 2.3e-92;

Matches 295; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKWTDLNHLFPISGLYKLNKSVQAQITLSILAMISTSLITAIL 60

DB 1 MSKNKQRTAKTLEKWTDLNHLFPISGLYKLNKSVQAQITLSILAMISTSLITAIL 60

QY 61 FIASANKVLTALIQDASQIKNTTPTLTQPOLGISFNSLSEITTSQTTILASTTP 120

DB 61 FIASANKVLTALIQDASQIKNTTPTLTQPOLGISFNSLSEITTSQTTILASTTP 120

QY 121 GYKSNLOPTVTKNTTPTTQPSKPTTKQKQKPNKPNNDHFVEVNFVPCISCSNP 180

DB 121 GYKSNLOPTVTKNTTPTTQPSKPTTKQKQKPNKPNNDHFVEVNFVPCISCSNP 180

QY 181 TCWAICKRIPNKKPKKTKTTTKPTTKPTTKKQKQKPTTKPKKPTTKPTTKPTTK 240

DB 181 TCWAICKRIPNKKPKKTKTTTKPTTKPTTKKQKQKPTTKPKKPTTKPTTKPTTK 240

QY 241 TMTITTLTNNNTGPKLTQSMTFHTSSSEGNLSPSQVSTTSEHPSPSPSPNTTQ 298

DB 241 TMTITTLTNNNTGPKLTQSMTFHTSSSEGNLSPSQVSTTSEHPSPSPSPNTTQ 298

##### RESULT 2

## MGZ

major surface glycoprotein G - human respiratory syncytial virus  
 C/Species: human respiratory syncytial virus  
 C/Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 24-Sep-1999  
 C/Accession: A94048; A93599; A04039  
 R/Wertz, G.W.; Collins, P.L.; Huang, Y.; Gruber, C.; Levine, S.; Ball, L.A.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 4075-4079, 1985  
 A/Title: Nucleotide sequence of the G protein gene of human respiratory syncytial virus  
 A/Reference number: A94048; MUID:85216636; PMID:3858865  
 A/Molecule type: mRNA  
 A/Residues: 1-298 <WER>  
 A/Cross-references: GB:M11486; GB:K01459; GB:K02719; GB:K03348; GB:M11217; GB:7.1; PID:g333932  
 A/Note: residues 207-298 are identical with residues 376-467 of the nucleocapsid protein  
 R/Satake, M.; Colligan, J.E.; Elango, N.; Norrby, E.; Venkatesan, S.  
 Nucleic Acids Res. 13, 7795-7812, 1985  
 A/Title: Respiratory syncytial virus envelope glycoprotein (G) has a novel structure.  
 A/Reference number: A93599; MUID:86067198; PMID:4069997  
 A/Accession: A93599  
 A/Molecule type: mRNA  
 A/Residues: 1-298 <SAT>  
 A/Cross-references: GB:K03149; MUID:g60997; PIDN:CAA26928.1; PID:g60998  
 C/Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C/Keywords: glycoprotein; transmembrane protein  
 F/38-66/Domain: transmembrane #status predicted <TNM>  
 F/85,135,237,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.8%; Score 1418; DB 1; Length 298;  
 Best Local Similarity 93.3%; Pred. No. 4.2e-85;  
 Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
 Qy 1 MSKNKDORTAKTLERKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
 Db 1 MSKNKDORTAKTLERKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
 Qy 61 FIASANHKVTLTIIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSOITTLASTTP 120  
 Db 61 FIASANHKVTLTIIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSOITTLASTTP 120  
 Qy 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPSCISNNP 180  
 Db 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPSCISNNP 180  
 Qy 181 TCWAICKRIIPNKKPGKTTTKPTTKPTTKKDLKPKQTTKPKVEPTTKTEPTINTTK 240  
 Db 181 TCWAICKRIIPNKKPGKTTTKPTTKPTTKKDLKPKQTTKPKVEPTTKTEPTINTTK 240  
 Qy 241 TNITTTLLTNTTGNPKLTSQMETFHTSSGNSLSPSQVSTTSEHPSQSPSPNTTR 298  
 Db 241 TNITTTLLTNTTGNPKLTSQMETFHTSSGNSLSPSQVSTTSEHPSQSPSPNTTR 298

## RESULT 3

JQ1205  
 attachment protein - human respiratory syncytial virus (strain RSB1734)  
 N/Alternate names: G protein  
 C/Species: human respiratory syncytial virus  
 C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C/Accession: JQ1205  
 R/Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A/Title: Identification of variable domains of the attachment (G) protein of subgroup A  
 A/Reference number: JQ1204; MUID:91374005; PMID:1895054  
 A/Accession: JQ1205  
 A/Molecule type: mRNA  
 A/Residues: 1-297 <CAN>  
 C/Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
 children and adults.  
 C/Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C/Keywords: glycoprotein; transmembrane protein  
 F/85,103,135,237,251,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.0%; Score 1343; DB 2; Length 297;  
 Best Local Similarity 88.6%; Pred. No. 3.1e-80;  
 Matches 263; Conservative 11; Mismatches 23; Indels 0; Gaps 0;  
 Qy 1 MSKNKDORTAKTLERKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
 Db 1 MSKNKDORTAKTLERKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
 Qy 61 FIASANHKVTLTIIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSOITTLASTTP 120  
 Db 61 FIASANHKVTLTIIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSOITTLASTTP 120  
 Qy 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPSCISNNP 180  
 Db 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPSCISNNP 180  
 Qy 181 TCWAICKRIIPNKKPGKTTTKPTTKPTTKKDLKPKQTTKPKVEPTTKTEPTINTTK 240  
 Db 181 TCWAICKRIIPNKKPGKTTTKPTTKPTTKKDLKPKQTTKPKVEPTTKTEPTINTTK 240  
 Qy 241 TNITTTLLTNTTGNPKLTSQMETFHTSSGNSLSPSQVSTTSEHPSQSPSPNTTR 297  
 Db 241 TNITTTLLTNTTGNPKLTSQMETFHTSSGNSLSPSQVSTTSEHPSQSPSPNTTR 297

## RESULT 4

JQ1208  
 attachment protein - human respiratory syncytial virus (strain RSB6256)  
 N/Alternate names: G protein  
 C/Species: human respiratory syncytial virus  
 C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C/Accession: JQ1208  
 R/Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A/Title: Identification of variable domains of the attachment (G) protein of subgroup A  
 A/Reference number: JQ1204; MUID:91374005; PMID:1895054  
 A/Accession: JQ1208  
 A/Molecule type: mRNA  
 A/Residues: 1-297 <CAN>  
 C/Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
 children and adults.  
 C/Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C/Keywords: glycoprotein; transmembrane protein  
 F/85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.0%; Score 1343; DB 2; Length 297;  
 Best Local Similarity 88.5%; Pred. No. 3.1e-80;  
 Matches 262; Conservative 8; Mismatches 26; Indels 0; Gaps 0;  
 Qy 1 MSKNKDORTAKTLERKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
 Db 1 MSKNKDORTAKTLERKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
 Qy 61 FIASANHKVTLTIIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSOITTLASTTP 120  
 Db 61 FIASANHKVTLTIIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSOITTLASTTP 120  
 Qy 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPSCISNNP 180  
 Db 121 SAESTQSTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPSCISNNP 180  
 Qy 181 TCWAICKRIIPNKKPGKTTTKPTTKPTTKKDLKPKQTTKPKVEPTTKTEPTINTTK 240  
 Db 181 TCWAICKRIIPNKKPGKTTTKPTTKPTTKKDLKPKQTTKPKVEPTTKTEPTINTTK 240  
 Qy 241 TNITTTLLTNTTGNPKLTSQMETFHTSSGNSLSPSQVSTTSEHPSQSPSPNTTR 296  
 Db 241 TNITTTLLTNTTGNPKLTSQMETFHTSSGNSLSPSQVSTTSEHPSQSPSPNTTR 296

## RESULT 5

JC5680

G protein - Human respiratory syncytial virus  
 C:Species: human respiratory syncytial virus  
 C>Date: 11-Nov-1997 #sequence\_revision 11-Nov-1997 #text\_change 26-Feb-1998  
 C:Accession: J05680  
 R:Geng, X.; Wang, Z.; Qian, Y.; Zhu, R.; Deng, J.; Du, J.; Zhu, Z.  
 Chinese J. Virol. 12, 317-322, 1996  
 A:Title: Molecular analysis of G protein gene of a respiratory syncytial virus strain is  
 A:Reference number: J05680  
 A:Accession: J05680  
 A:Molecule type: mRNA  
 A:Residues: 1-298 <GEN>  
 A:Experimental source: strain B79  
 A:Note: the authors translated the codon TTT for residue 165 and 170 as Glu, TTC for res  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 F:1-38/Domain: intracellular #status predicted <INT>  
 F:39-66/Domain: transmembrane #status predicted <TM>  
 F:67-298/Domain: extracellular #status predicted <EXC>

Query Match 86.3%; Score 1333; DB 2; Length 298;  
 Best Local Similarity 87.9%; Pred. No. 1.4e-79;  
 Matches 262; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60  
 DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60

QY 61 FIASANKHVLTITAIIDQATSOIKNTTPTTYLTQDPLGIFSLSNLSEITSTQTTILASTTP 120  
 DB 61 FIASANKHVLTITAIIDQATNOIKNTTPTTYLTQDPLGIFSLSNLSEITSTQTTILASTTP 120

QY 121 GVKSNLQPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
 DB 121 SAESTPQSTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180

QY 181 TCWAICKRIPNKKPGKKTTPKTKPTKTKKQKPKKPNKPNNDHFEVFNVPVCSICSNPP 240  
 DB 181 TCWAICKRIPNKKPGKKTTPKTKPTKTKKQKPKKPNKPNNDHFEVFNVPVCSICSNPP 240

QY 241 TNITTTLLTNTTGNPKLTSMETPHSTSSGNSLSPSOVSTTSSEHPSQSPSPNNT 296  
 DB 241 TNIRTTLLTNTTGNPKLTSMETPHSTSSGNSLSPSOVSTTSSEHPSQSPSPNNT 296

RESULT 6  
 J01209  
 attachment protein - human respiratory syncytial virus (strain RSB6614)  
 N:Alternate names: G protein  
 C:Species: human respiratory syncytial virus  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C:Accession: J01209  
 R:Can, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
 A:Reference number: J01209; MUID:91374005; PMID:1895054  
 A:Accession: J01209  
 A:Molecule type: mRNA  
 A:Residues: 1-297 <CAN>  
 C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract in  
 children and adults.  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C:Keywords: glycoprotein; transmembrane protein  
 F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 86.0%; Score 1328; DB 2; Length 297;  
 Best Local Similarity 87.8%; Pred. No. 2.9e-79;  
 Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60  
 DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60

QY 61 FIASANKHVLTITAIIDQATSOIKNTTPTTYLTQDPLGIFSLSNLSEITSTQTTILASTTP 120  
 DB 61 FIASANKHVLTITAIIDQATNOIKNTTPTTYLTQDPLGIFSLSNLSEITSTQTTILASTTP 120

QY 121 GVKSNLQPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
 DB 121 SAESTPQSTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180

QY 181 TCWAICKRIPNKKPGKKTTPKTKPTKTKKQKPKKPNKPNNDHFEVFNVPVCSICSNPP 240  
 DB 181 TCWAICKRIPNKKPGKKTTPKTKPTKTKKQKPKKPNKPNNDHFEVFNVPVCSICSNPP 240

QY 241 TNITTTLLTNTTGNPKLTSMETPHSTSSGNSLSPSOVSTTSSEHPSQSPSPNNT 296  
 DB 241 TNIRTTLLTNTTGNPKLTSMETPHSTSSGNSLSPSOVSTTSSEHPSQSPSPNNT 296

RESULT 7  
 J01206  
 attachment protein - human respiratory syncytial virus (strain RSB5857)  
 N:Alternate names: G protein  
 C:Species: human respiratory syncytial virus  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C:Accession: J01206  
 R:Can, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A:Title: Identification of variable domains of the attachment (G) protein of subgroup  
 A:Reference number: J01204; MUID:91374005; PMID:1895054  
 A:Accession: J01206  
 A:Molecule type: mRNA  
 A:Residues: 1-297 <CAN>  
 C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract  
 children and adults.  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C:Keywords: glycoprotein; transmembrane protein  
 F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 85.8%; Score 1324; DB 2; Length 297;  
 Best Local Similarity 87.8%; Pred. No. 5.3e-79;  
 Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60  
 DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60

QY 61 FIASANKHVLTITAIIDQATSOIKNTTPTTYLTQDPLGIFSLSNLSEITSTQTTILASTTP 120  
 DB 61 FIASANKHVLTITAIIDQATNOIKNTTPTTYLTQDPLGIFSLSNLSEITSTQTTILASTTP 120

QY 121 GVKSNLQPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
 DB 121 SAESTPQSTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180

QY 181 TCWAICKRIPNKKPGKKTTPKTKPTKTKKQKPKKPNKPNNDHFEVFNVPVCSICSNPP 240  
 DB 181 TCWAICKRIPNKKPGKKTTPKTKPTKTKKQKPKKPNKPNNDHFEVFNVPVCSICSNPP 240

QY 241 TNITTTLLTNTTGNPKLTSMETPHSTSSGNSLSPSOVSTTSSEHPSQSPSPNNT 296  
 DB 241 TNIRTTLLTNTTGNPKLTSMETPHSTSSGNSLSPSOVSTTSSEHPSQSPSPNNT 296

RESULT 8  
 J01207  
 attachment protein - human respiratory syncytial virus (strain RSB6190)  
 N:Alternate names: G protein  
 C:Species: human respiratory syncytial virus  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C:Accession: J01207  
 R:Can, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A:Title: Identification of variable domains of the attachment (G) protein of subgroup  
 A:Reference number: J01204; MUID:91374005; PMID:1895054  
 A:Accession: J01207  
 A:Molecule type: mRNA

Query Match 86.0%; Score 1328; DB 2; Length 297;  
 Best Local Similarity 87.8%; Pred. No. 2.9e-79;  
 Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60  
 DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60

QY 61 FIASANKHVLTITAIIDQATSOIKNTTPTTYLTQDPLGIFSLSNLSEITSTQTTILASTTP 120  
 DB 61 FIASANKHVLTITAIIDQATNOIKNTTPTTYLTQDPLGIFSLSNLSEITSTQTTILASTTP 120



Best Local Similarity 54.9%; Pred. No. 1.7e-42;  
Matches 158; Conservative 35; Mismatches 93; Indels 2; Gaps 2;

QY 1 MSKNKQRTAKTLEKTDWTLNHLIFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
Db 1 MSKHQRTAKTLEKTDWTLNHLIVISSCLYRLNLSIAQIALSVLAMIISTLSIIITAI 60  
QY 61 FIASANKHVLTITAIIDQATSOIKNTPTTYLTODPOLGIFSLSNLSITSTTTILASTTP 120  
Db 61 FIISANKHVLTITVQTIKNHTEKNTISTYLTODPOLGIFSLSNLSITSTTTILASTTP 120  
QY 121 GVKSNLOPTTKNTTITOTOPSKPTTKORQKNPKPNNDPHEFVFNVPVCSICSNPP 180  
Db 121 NTKSETHHTAQTKRITITSTQTNKPSKSRNPPKPKDDVHEFVFNVPVCSICSNQ 180  
QY 181 TCMAICKRIIPNKKPKKTKPTTKPTKTT-KDGLKPOQTKPKVEPTTKPTKEPTINTT 239  
Db 181 LKSIKTIIPNKKPKKTKPTTKPTTKTKNREDPKTPAKMPKKEIITNPAPKPKTKTT 240  
QY 240 KTNITITLLTNNTGNPKLTSOMETHSTSEGNLSPSQVSTTSEHPS 287  
Db 241 ERDTSISQSTVLDITTPKTYIQOQSLHSTTSENTPSTQIPTASE-PS 287

## RESULT 12

VNZ

nucleocapsid protein (version 2) - human respiratory syncytial virus

C:Species: human respiratory syncytial virus

C>Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 16-Jul-1999

C:Accession: A04026

R:Elango, N.; Venkatesan, S.

Nucleic Acids Res 11, 5941-5951, 1983

A:Title: Amino acid sequence of respiratory syncytial virus capsid protein.

A:Reference number: A04026; MUID:83299261; PMID:6310521

A:Accession: A04026

A:Molecule type: mRNA

A:Residues: 1-467 <EUA>

A:Cross-references: GB:X00001; NID:G61215; PIDN:CRA24906.1; PID:G61216

C:Genetics:

A:Gene: N

C:Superfamily: respiratory syncytial virus nucleocapsid protein

C:Keywords: nucleocapsid

Query Match 27.5%; Score 425; DB 1; Length 467;  
Best Local Similarity 90.2%; Pred. No. 1.7e-20;  
Matches 83; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 207 TFKTKKDLKPKQTKPKVEPTTKPTKEPTINTTKNITITLLTNNTGNPKLTSQMETPH 266  
Db 376 TLKTKKDKPKQTKPKVEPTTKPTKEPTINTTKNITITLLTNNTGNPKLTSQMETPH 435

QY 267 STSSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298

Db 436 STSSEGNLSPSQVSTTSEHPSQSPSPNTTQ 467

## RESULT 13

PQ0768

glycoprotein G - bovine respiratory syncytial virus (isolate FS-1) (fragment)

C:Species: bovine respiratory syncytial virus

C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999

C:Accession: PQ0768

R:Malipeddi, S.K.; Samal, S.K.

J. Gen. Virol. 74, 2001-2004, 1993

A:Title: Sequence variability of the glycoprotein gene of bovine respiratory syncytial

A:Reference number: J02284; MUID:93389461; PMID:8376974

A:Accession: PQ0768

A:Molecule type: mRNA

A:Residues: 1-250 <MAL>

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:1-31/Domain: intracellular #status predicted <INT>

F:32-59/Domain: transmembrane #status predicted <TM>

F:60-250/Domain: extracellular #status predicted <EXT>  
F:78,156,226,241,244/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.4%; Score 299; DB 2; Length 250;  
Best Local Similarity 30.3%; Pred. No. 1.3e-12;  
Matches 91; Conservative 35; Mismatches 104; Indels 70; Gaps 9;

QY 11 KLEKNTWTLNHLIFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAIIFIASANKHV 70  
Db 4 KTLKRAWKASYFIVGLSCLYKFNLSLVQVATLSTLAMIITLSIVITAIIVISGNAKAK 63  
QY 71 LTTALIQDQATSOIKNTPTTYLTODPOLGIFSLSNLSITSTTTILASTTPGVKSNLOPTT 130  
Db 64 PTKPTTIQOTQOPQNHSTSPFFIENY-----KSTHTSIQSTTLLSOLINIDTTR 111  
QY 131 VKTKNTTITOTQPSK-----PTTKQ-----RONKPPKNFNDPHEFVFNVPVCSICSN 178  
Db 112 GTTYGHSTDEIQSKIKSQSTLPTTRKPPINPSSNPENHODHNSQTLPIVPCSTCEG 171  
QY 179 NPTCHAIKRIIPNKKPKKTKPTTKPTTKKDLKPOQTKPKVEPTTKPTKEPTINTT 238  
Db 172 NLACLSLCQVGPGR-----APSRAPTI-TLAKTKPKTK-----KP----- 207  
QY 239 TKTNITITLLTNNTGNPKLTSQMETHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298  
Db 208 ----IKTTI-----HRTSPKALQPKN-NTAAPQOGILSSPENTNQ 245

## RESULT 14

MGNZBR

major surface glycoprotein G - bovine respiratory syncytial virus (strain 391-2)

N:Alternate names: attachment glycoprotein G

C:Species: bovine respiratory syncytial virus

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999

C:Accession: A36408

R:Lerch, R.A.; Anderson, K.; Wertz, G.W.

J. Virol. 64, 5559-5569, 1990

A:Title: Nucleotide sequence analysis and expression from recombinant vectors demonst

Y syncytial virus.

A:Reference number: A36408; MUID:91012801; PMID:2214024

A:Accession: A36408

A:Molecule type: mRNA

A:Residues: 1-257 <LER>

A:Cross-references: GB:M8307; NID:G210830; PIDN:AAA42810.1; PID:G210831

C:Genetics:

A:Gene: G

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:45-62/Domain: transmembrane #status predicted <TM>

F:73,85,127,149,233,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.4%; Score 299; DB 1; Length 257;  
Best Local Similarity 31.6%; Pred. No. 1.3e-12;  
Matches 92; Conservative 33; Mismatches 120; Indels 46; Gaps 8;

QY 1 MSKNKQRTAKTLEKTDWTLNHLIFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
Db 1 MSNHTHLKFKTLKRAWKASYFIVGLSCLYKFNLSLVQVATLSTLAMIITLSIVITAI 60  
QY 61 FIASANKHVLTITAIIDQATSOIKNTPTTYLTODPOLGIFSLSNLSITSTTTILASTTP 120  
Db 61 YISVGNARAKPTKPTTIQOTQOPQNHSTSPFFIENY-----KSTHTSIQSTTLL 108  
QY 121 GVKSNLOPTTKNTTITOTQPSK---PTTKQKNPKPNNDP-----HPEVF 168  
Db 109 SOLLNIDITRGITYGHSTNEQNKRKIQOSTLPATRKPPINPSSIPENHODHNNQTL 168  
QY 169 NFVPCSICSNPTCWAICKRIIPNKKPKKTKPTTKPTTKKDLKPOQTKPKVEPTT 228  
Db 169 PVVPCSTCEGNLACLSLCH-----IETERPAPRTI-TLAKTKPKTK----- 212  
QY 229 KPTKEPTINTTKNITITLLTNNTGNPKLTSQMETHSTSEGNLSPSQV 279

Db 213 KPTK--TTIHRSTPETKLPKNNATP-----QQGILSSTEHHNQSTTQI 257

## RESULT 15

JQ2284

glycoprotein G - bovine respiratory syncytial virus (isolate A51908)

C:Species: bovine respiratory syncytial virus

C&gt;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999

C:Accession: JQ2284

R:Mailipedi, S.K.; Samal, S.K.

J. Gen. Virol. 74, 2001-2004, 1993

A:Title: Sequence variability of the glycoprotein gene of bovine respiratory syncytial v

A:Reference number: JQ2284; MUID:93389461; PMID:8376974

A:Accession: JQ2284

A:Molecule type: mRNA

A:Residues: 1-263 &lt;MAL&gt;

A:Experimental source: isolate A51908

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:1-38/Domain: intracellular #status predicted &lt;CYT&gt;

F:39-66/Domain: transmembrane #status predicted &lt;TM&gt;

F:67-263/Domain: extracellular #status predicted &lt;EXT&gt;

F:127,163,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

|    | Query Match           | 19.1%  | Score 295.5;       | DB 2; | Length 263;                        |
|----|-----------------------|--|--------------------|-------|------------------------------------|
|    | Best Local Similarity | 31.7%  | Pred. No. 2.3e-12; |       |                                    |
|    | Matches               | 86;  | Conservative       | 35;   | Mismatches 123; Indels 27; Gaps 7; |
| Qy | 1                     | MSKNKQRTATLEKTDNLHLLFISSGLYKLNKSVQAQITLSILAMISTSLITAI      | 60                 |       |                                    |
| Db | 1                     | MSNHTHPKFKTKRAMKASKYFIVGLSCLYFNLSIVQALTSLAMITLSLVITAI      | 60                 |       |                                    |
| Qy | 61                    | FIASANKHVTLTALIQDATSQIKNTTPTYLTDPPQLGISFNSLSEITSTTILASTTP  | 120                |       |                                    |
| Db | 61                    | YISVGNKAKETSKPTTQQTQOPQNHPTLLPTEHNHKSHTSTSTSTLSQPPNI--DTTS | 118                |       |                                    |
| Qy | 121                   | GVK-----SNLQPTTVTKNT--TTQTPSKPTTKQRONKPPKPNNDHFVEFNVFPCS   | 174                |       |                                    |
| Db | 119                   | GTTVGHPINRQNRKIKSQSTPLATRLPINPL-----ESNPENHQDHNSQTLPHVPCS  | 174                |       |                                    |
| Qy | 175                   | ICSNNPTCAIC----KRIPN-----KPGKKTTKTKPKTKTKKDLKPQTKP         | 222                |       |                                    |
| Db | 175                   | TCEGNPACSPICQIGLERAPSRAPTITLKAKPKTKTKTKTKTIYHRTSPEAKLQTKN  | 234                |       |                                    |
| Qy | 223                   | KEVPTTKPTERTPTINTTKTNTITLTNTNTT                            | 253                |       |                                    |
| Db | 235                   | TATPQQGILSP---EHGTNQSTTQISQHTS                             | 262                |       |                                    |

Search completed: October 29, 2003, 17:41:49

Job time : 31.6755 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:41:00 ; Search time 52.2906 Seconds  
(without alignments) 975.349 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDQRTAKTLEKTWDTL.....VSTSEHPSQPSSPPNTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 642050 seqs. 171146064 residues

Total number of hits satisfying chosen parameters: 642050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

### Listing first 45 summaries

```

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/FCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/FCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query |      | Length | DB                  | ID                 | Description |
|------------|-------|-------|------|--------|---------------------|--------------------|-------------|
|            |       | Match | %    |        |                     |                    |             |
| 1          | 193   | 12.5  | 5179 | 9      | US-09-922-217-1068  | Sequence 1068, Ap  |             |
| 2          | 193   | 12.5  | 5179 | 10     | US-09-833-263-1068  | Sequence 1068, Ap  |             |
| 3          | 193   | 12.5  | 5179 | 14     | US-10-025-380-1068  | Sequence 1068, Ap  |             |
| 4          | 177.5 | 11.5  | 800  | 12     | US-10-029-386-32198 | Sequence 32198, A  |             |
| 5          | 169.5 | 11.0  | 1367 | 10     | US-09-801-368-108   | Sequence 108, App  |             |
| 6          | 169   | 10.9  | 36   | 12     | US-10-351-641-871   | Sequence 871, App  |             |
| 7          | 162.5 | 10.5  | 288  | 9      | US-09-216-393-341   | Sequence 341, App  |             |
| 8          | 162.5 | 10.5  | 288  | 9      | US-09-216-393-344   | Sequence 344, App  |             |
| 9          | 162.5 | 10.5  | 288  | 12     | US-10-321-856-341   | Sequence 341, App  |             |
| 10         | 162.5 | 10.5  | 288  | 12     | US-10-321-856-344   | Sequence 344, App  |             |
| 11         | 162   | 10.5  | 941  | 14     | US-10-124-557-14    | Sequence 14, Appl  |             |
| 12         | 162   | 10.5  | 1022 | 14     | US-10-124-557-84    | Sequence 84, Appl  |             |
| 13         | 162   | 10.5  | 1038 | 14     | US-10-124-557-74    | Sequence 74, Appl  |             |
| 14         | 162   | 10.5  | 1049 | 14     | US-10-124-557-58    | Sequence 58, Appl  |             |
| 15         | 162   | 10.5  | 1140 | 14     | US-10-124-557-104   | Sequence 104, Appl |             |

## ALIGNMENTS

## RESULT 1

```

US-09-922-217-1068
; Sequence 1068, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagner, Madeline Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: OF COLON CANCER AND METHODS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-922-217-1068

```

Query Match 12.5%; Score 193; DB 9; Length 5179;  
Best Local Similarity 30.6%; Pred. No. 2.3e-06;  
Matches 77. Conservative 23. Mismatches 114. Indels 3

```
QY      71 LTTALIQDA--TSQIKNTTP-----TYLTQDPQLGIFSNLSEIT--SQTTILASTPG 121
       :|||   |:::|||   |   |   |   |   |   |   |   |   |   |   |   |
Db     1504 MTTPIPTPRASTTLLPPTTSPSPPTTTTTTTPPTTTPSPPTTPIPTPTSTTTLPPPTPS 1563
```

|    |      |   |  |                           |      |
|----|------|---|--|---------------------------|------|
| Qy | 122  | ----  | VKSNLOPTTVTKNNTTT-----                     | QTQPSKPITTKQRONKPPNKPNDFH | 164  |
| Dd | 1564 | PPTTTTTTTPPPTTSPSPPITTTTTPPPTTTPSPPITTTTTPPPTTTPSPPIT   | :  | :                         |      |
| Qy | 165  | FYFNFVPCSGICNNPFCWALCKRIPNKKGSKITTKTAKPIKPTTKOLKQPOTPKB | :  | :                         | 224  |
| Dd | 1624 | TPI--TWPTSTTLDPPT-----                                  | TTSPSPPTTTTTTTPPPTTTPSPPITTTTPSPPITTTTTTTP |                           | 1675 |
| Qy | 225  | VPTTKTEEPITNTKITNTLTLLTNNTNGPKLTSQMETHSTSSEGNLSPOSVTSS  | :  | :                         | 284  |
| Dd | 1676 | PPITTPSS--PIITTPSPPTTTWTTPPTTP--                        | SSPITTTTPSSITTTSPSPPITMTTP                 |                           | 1731 |
| Qy | 285  | HPSQPSPPNPT   | 296  |                           |      |
| Dd | 1732 | SPTTTPSPPTT   | 1743                                       |                           |      |

```

RESULT 2
US-09-833-263-1068
; Sequence 1068, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-833-263-1068

```

|                       |        |                    |                 |                    |
|-----------------------|--------|--------------------|-----------------|--------------------|
| Query Match           | 12.5%; | Score 193;         | DB 10;          | Length 5179;       |
| Best local Similarity | 30.6%; | Pred. No. 2.3e-06; |                 |                    |
| Matches               | 77;    | Conservative 23;   | Mismatches 114; | Indels 38; Gaps 9; |

  

|    |      |  |      |
|----|------|--|------|
| QY | 71   | LTTAIIQDA--TSQIKNTP-----TVLTQDPQLGIGSFNSLSEIT--SQTTILASTTGG  | 121  |
| Db | 1504 | MTTPEITPPASTIILPPTTSSPPTTTTTTPPTTSPPTTPTTPTSTIPLPPTTTPS      | 1563 |
| QY | 122  | -----VKSNLOPTVVKTKNTT-----QTQPSKPTTKQONKPKPNKNDFH            | 164  |
| Db | 1564 | PPPTTTTTPPTTTPSPPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPSPPT    | 1623 |
| QY | 165  | FEVFNFPVPCISGNPNPCWAI CKRI PNKKPGKKTTKTKPKTEPKTKKDLKPOTTRPKE | 224  |
| Db | 1624 | TPI--TPETSIILPPT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP     | 1675 |
| QY | 225  | VPTTKPREPTINTKTNITTTTLNNNTGNPKLTSQMETHFSTPSSEGNLSPQSVSTTSE   | 284  |
| Db | 1676 | PPPTTPES--PITTPSPPTTMTTPPTTTP--SSPITTTTPSTSTTTPSPPTTMTPT     | 1731 |

RESULT 3  
US-10-025-380-1068  
; Sequence 1068, Application US/10025380  
; Publication No. US20020182191A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.

```

; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-025-380-1068

```

|                       |        |  |                 |                     |
|-----------------------|--------|--|-----------------|---------------------|
| Query Match           | 12.5%; | Score 193;   | DB 14;          | Length 5179;        |
| Best Local Similarity | 30.6%; | Pred. No. 2.3e-06;   |                 |                     |
| Matches               | 77;    | Conservative 23;   | Mismatches 114; | Indels 38; Gaps 19; |
| QY                    | 71     | LITAILQDA--TSQIKNTTP-----TYLQDPQIGISFSLSEIT--SQTTTILASTTPG | 121             |                     |
| Db                    | 1504   | MTTFITPPASTTLLPTTTPSPPTTTTTPPTTTPSPPTTITTPSTSTTLLPTTTP     | 1563            |                     |
| QY                    | 122    | ----VKSNLQPTTVKKNKNTTT-----QTQSPKPTTKQKQKPPKPKNNDFH        | 164             |                     |
| Db                    | 1564   | PPPTTTTPPTTTPSPPTTTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPSPPTT     | 1623            |                     |
| QY                    | 165    | FEVFNFPVCSICNNPPTCMAICKRIKPKKGKTKTKTKPKTKTKTKDLKPKQTTPKPE  | 224             |                     |
| Db                    | 1624   | TPI--TPPTSTTLLPT-----TTPSPPTTTPPTTTPSPPTTTPSPPTTTPPTTTP    | 1675            |                     |
| QY                    | 225    | VPTTKTPEETINTKTIHTLLTNNNTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSE   | 284             |                     |
| Db                    | 1676   | PPPTTSS--PIITTSPPPTTMTTSPITTP--SPITTTTTTSPSTTTPSPPTTMTTP   | 1731            |                     |
| QY                    | 285    | HPSQPSPPNTT  | 296             |                     |
| Db                    | 1732   | SPTTTSPPTTT  | 1743            |                     |

```

RESULT 4
US-10-029-386-32198
; Sequence 32198, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32198
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007663.28
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2

```





Db 76 AALSTDAIDRVSDVLSLLVIREAAQAFDLGLRLITDIASGIGEGAWALMGEEAAFI 135  
Qy 84 -----XNTPYLAQDPOLGIFSGLSEITSSQTTILASTTPGVKSNLQPTTVTKTKN 135  
Db 136 RPRRSKGGKTTT-----TSSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTP 181  
Qy 136 TTTTQTPSKPTTKQKQKPNKPNDFHFEVFNFCSCNNPTCWAICKIPNKKPG 195  
Db 182 TTTTTTTTTTT----- 194  
Qy 196 KKTTPKTKPTTKKOLKQTPKQTPKQTPKQTPKQTPKQTPKQTPKQTPKQTPKQTP 296  
Db 195 -TTE 249  
Qy 256 PKLTSQMETFHSSEGNLSPSQVSTTSEHPSQSPSPNTT 296  
Db 250 PTTT-----TTTPTTTTSTTTTTTTTTTTTTTTTTTTTTTTTTTTT 284  
  
RESULT 11  
US-10-124-557-14  
; Sequence 14, Application US/10124557  
; Publication No. US20020137894A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Katherine  
; Clark, Stephen C.  
; Jacobs, Kenneth  
; Hewick, Rodney M.  
; Gesner, Thomas G.  
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; City: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/124,557  
; FILING DATE: 16-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/643,502  
; FILING DATE: 18-JAN-1991  
; APPLICATION NUMBER: US 07/546,114  
; FILING DATE: 29-JUN-1990  
; APPLICATION NUMBER: US 07/457,196  
; FILING DATE: 29-DEC-1989  
; APPLICATION NUMBER: US 07/390,901  
; FILING DATE: 08-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cserr, Luann  
; REGISTRATION NUMBER: 31,822  
; REFERENCE/DOCKET NUMBER: GI 5190  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)876-1170  
; TELEFAX: (617)876-5851  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 941 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-10-124-557-14

Query Match 10.5%; Score 162; DB 14; Length 941;  
Best Local Similarity 24.6%; Pred. No. 9e-05;  
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;  
  
Qy 72 TTAIQDATSOKNTTPYLVQDQLGIFSGLSEITSSQTTILASTTPGVKSNLQPTTV 131  
Db 135 TSKVLAKTPKAEITTK-----GPAITTPKPTTPKPEASTTP---KEPTPTTI 172  
Qy 132 K-----TKNTTTTQTPSKPTTKQKQKPNKPNDFHFEVFNFCSCSN 178  
Db 173 KSAPITPKPEAPTTTKSAPITPKPEAPTTTKPEAPTTTKPE-----APTITKEP 221  
Qy 179 NPTCWAICKIPN-----KPKGKTTTKP-----TKKPTFTTK 212  
Db 222 APTITKSAPITPKPEAPTTPKKPAITPKPEAPTTTKPEAPTTTKPEAPTTTKPEAPTTK 281  
Qy 213 KDL-----KPOITPKPEVPTTKP-----TEPTINTTKINITTTILTNNTGNPKLTS 260  
Db 282 EPAPTAPKKPAITPKPEAPTTTKPEAPTTTKPEAPTTTKPEAPTTTKPEAPTTTKPEAPTT 341  
Qy 261 QMETFHSSEGNLSPSQVSTTSEHPSQSPSPNTT 296  
Db 342 TTKSAPITPKPE-----PSPTTKPEAPTTTKPEAPTT 373  
  
RESULT 12  
US-10-124-557-84  
; Sequence 84, Application US/10124557  
; Publication No. US20020137894A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Katherine  
; Clark, Stephen C.  
; Jacobs, Kenneth  
; Hewick, Rodney M.  
; Gesner, Thomas G.  
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; City: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/124,557  
; FILING DATE: 16-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/643,502  
; FILING DATE: 18-JAN-1991  
; APPLICATION NUMBER: US 07/546,114  
; FILING DATE: 29-JUN-1990  
; APPLICATION NUMBER: US 07/457,196  
; FILING DATE: 29-DEC-1989  
; APPLICATION NUMBER: US 07/390,901  
; FILING DATE: 08-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cserr, Luann  
; REGISTRATION NUMBER: 31,822  
; REFERENCE/DOCKET NUMBER: GI 5190  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)876-1170  
; TELEFAX: (617)876-5851  
; INFORMATION FOR SEQ ID NO: 84:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1022 amino acids  
; TYPE: amino acid

```
;
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-10-124-557-84

Query Match      10.5%; Score 162; DB 14; Length 1022;
Best Local Similarity 24.6%; Pred. No. 0.0001;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 72 TTAIIQDASQIKNTTPTLYLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNLQPTTV 131
Db 206 TSKVLAKPTPKAETTK-----GPALTTPKEPTTPPKBPASTTP---KEPTPTTI 253
QY 132 K-----TKNTTTTQTPSKPTTKQKQKPNKPNNDFFHEVFNFVPCSCSN 178
Db 254 KSAPTTPKEPAPTTTKSAPTTTPKEPAPTTTPKEPAPTTTPKEP-----APTTPKEP 302
QY 179 NPTCWAICKRIPN-----KKPKGKTTTKP-----TKKPTFKTK 212
Db 303 APTTKSAPTTTPKEPAPTTTPKAPTTTPKEPAPTTTPKEPAPTTTPKEPAPTTTPK 362
QY 213 KDL-----KPQTTKPKVPTTKP-----TEPTINTTKNTITLLTNNTGNPKLTS 260
Db 363 EPAPTAPKAPTTTPKEPAPTTTPKEPAPTTTPKPSPTTPKEPAPTTTKSAPTTTPKEPAPT 422
QY 261 QMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTT 296
Db 423 TTKSAPTTPKP-----PSPTTPKEPAPTTTPKEPAPT 454

RESULT 13
US-10-124-557-74
; Sequence 74, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
;           Clark, Stephen C.
;           Jacobs, Kenneth
;           Hewick, Rodney M.
;           Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserit, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
```

```
;
;   TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 74:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1038 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;     SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-124-557-74

Query Match      10.5%; Score 162; DB 14; Length 1038;
Best Local Similarity 24.6%; Pred. No. 0.0001;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 72 TTAIIQDASQIKNTTPTLYLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNLQPTTV 131
Db 190 TSKVLAKPTPKAETTK-----GPALTTPKEPTTPPKBPASTTP---KEPTPTTI 237
QY 132 K-----TKNTTTTQTPSKPTTKQKQKPNKPNNDFFHEVFNFVPCSCSN 178
Db 238 KSAPTTPKEPAPTTTKSAPTTTPKEPAPTTTPKEPAPTTTPKEP-----APTTPKEP 286
QY 179 NPTCWAICKRIPN-----KKPKGKTTTKP-----TKKPTFKTK 212
Db 287 APTTKSAPTTTPKEPAPTTTPKAPTTTPKEPAPTTTPKEPAPTTTPKEPAPTTTPK 346
QY 213 KDL-----KPQTTKPKVPTTKP-----TEPTINTTKNTITLLTNNTGNPKLTS 260
Db 347 EPAPTAPKAPTTTPKEPAPTTTPKEPAPTTTPKPSPTTPKEPAPTTTKSAPTTTPKEPAPT 406
QY 261 QMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTT 296
Db 407 TTKSAPTTPKP-----PSPTTPKEPAPTTTPKEPAPT 438

RESULT 14
US-10-124-557-58
; Sequence 58, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
;           Clark, Stephen C.
;           Jacobs, Kenneth
;           Hewick, Rodney M.
;           Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserit, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:36:05 ; Search time 16.3057 Seconds  
(without alignments)  
773.267 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDQRTAKTLEKTDYTL.....VSTSEHPSPSPPTTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/iaa/5A COMB.pdp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pdp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pdp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pdp.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pdp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 1544  | 100.0       | 298    | 2     | US-08-467-963C-8   |
| 2          | 1544  | 100.0       | 298    | 2     | US-08-838-189D-8   |
| 3          | 1544  | 100.0       | 298    | 3     | US-08-852-344D-8   |
| 4          | 1544  | 100.0       | 298    | 3     | US-08-344-639E-8   |
| 5          | 1544  | 100.0       | 298    | 3     | US-08-467-969A-8   |
| 6          | 1544  | 100.0       | 298    | 3     | US-08-467-961A-8   |
| 7          | 1544  | 100.0       | 298    | 3     | US-08-001-554A-8   |
| 8          | 901   | 58.4        | 681    | 6     | 5194595-19         |
| 9          | 295.5 | 19.1        | 263    | 5     | PCT-US91-08177-13  |
| 10         | 222   | 14.4        | 37     | 3     | US-08-793-792-12   |
| 11         | 193   | 12.5        | 32     | 3     | US-08-793-792-8    |
| 12         | 188   | 12.2        | 216    | 3     | US-08-928-361B-8   |
| 13         | 188   | 12.2        | 216    | 3     | US-08-928-361B-8   |
| 14         | 188   | 12.2        | 1837   | 4     | US-08-928-361B-5   |
| 15         | 188   | 12.2        | 1837   | 4     | US-08-588-995A-5   |
| 16         | 187.5 | 12.1        | 1721   | 3     | US-08-700-651-5    |
| 17         | 187.5 | 12.1        | 1721   | 3     | US-08-928-361B-6   |
| 18         | 187.5 | 12.1        | 1721   | 3     | US-08-588-995A-6   |
| 19         | 183.5 | 11.9        | 216    | 3     | US-08-928-361B-27  |
| 20         | 174   | 11.3        | 28     | 3     | US-08-793-792-4    |
| 21         | 173.5 | 11.2        | 249    | 3     | US-08-700-651-15   |
| 22         | 173.5 | 11.2        | 249    | 3     | US-08-928-361B-20  |
| 23         | 173.5 | 11.2        | 249    | 4     | US-08-588-995A-20  |
| 24         | 169   | 10.9        | 36     | 3     | US-09-082-279B-871 |
| 25         | 169   | 10.9        | 36     | 4     | US-09-315-304B-871 |
| 26         | 169   | 10.9        | 36     | 4     | US-09-834-784-871  |
| 27         | 164.5 | 10.7        | 2476   | 2     | US-08-276-967-2    |

28 162.5 10.5 288 4 US-09-216-393B-341 Sequence 341, App  
29 162.5 10.5 288 4 US-09-216-393B-344 Sequence 344, App  
30 162 10.5 941 4 US-07-757-022B-14 Sequence 14, Appl  
31 162 10.5 1022 4 US-07-757-022B-84 Sequence 84, Appl  
32 162 10.5 1038 4 US-07-757-022B-74 Sequence 74, Appl  
33 162 10.5 1049 4 US-07-757-022B-58 Sequence 58, Appl  
34 162 10.5 1140 4 US-07-757-022B-104 Sequence 104, App  
35 162 10.5 1270 4 US-07-757-022B-44 Sequence 44, Appl  
36 162 10.5 1311 4 US-07-757-022B-42 Sequence 42, Appl  
37 162 10.5 1313 4 US-07-757-022B-142 Sequence 142, App  
38 162 10.5 1314 4 US-07-757-022B-50 Sequence 50, Appl  
39 162 10.5 1320 4 US-07-757-022B-46 Sequence 46, Appl  
40 162 10.5 1320 4 US-07-757-022B-40 Sequence 60, Appl  
41 162 10.5 1354 4 US-07-757-022B-48 Sequence 48, Appl  
42 162 10.5 1361 4 US-07-757-022B-52 Sequence 40, Appl  
43 162 10.5 1363 4 US-07-757-022B-2 Sequence 52, Appl  
44 162 10.5 1404 4 US-07-757-022B-2 Sequence 2, Appl  
45 162 10.5 1404 4 US-07-757-022B-62 Sequence 62, Appl

#### ALIGNMENTS

##### RESULT 1

US-08-467-963C-8

; Sequence 8, Application US/08467963C

; Patent No. 5968776

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michel H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHYN, Mary E

; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES PROTECTION AGAINST

; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6th Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER: US/08/467.963C

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION NUMBER: US 08/838,189

; FILING DATE: 16-APR-1997

; APPLICATION NUMBER: US 08/001,554

; FILING DATE: 06-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9200117.1

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 298 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-467-963C-8

Query Match 100.0%; Score 1544; DB 2; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.1e-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
DB 1 MSKNKQORTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANEKVTLTAI IQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTOTTTILASTTP 120  
DB 61 FIASANEKVTLTAI IQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTOTTTILASTTP 120

QY 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQONKPNKPNNDHFEVFNFPVPCISCSNNP 180  
DB 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQONKPNKPNNDHFEVFNFPVPCISCSNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPTKKPTFKTKDLPQTKPKKEVPTTKPTTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKTTTKPTKKPTFKTKDLPQTKPKKEVPTTKPTTEPTINTTK 240

QY 241 TNITTTLLTNNNTGNPKLTSMETFHSTSEGNLSPSOVSTTSEHPSPQSPSPNTTRQ 298  
DB 241 TNITTTLLTNNNTGNPKLTSMETFHSTSEGNLSPSOVSTTSEHPSPQSPSPNTTRQ 298

## RESULT 2

US-08-838-189D-8  
; Sequence 8, Application US/08838189D  
; Patent No. 5998169  
; GENERAL INFORMATION:  
; APPLICANT: KLEIN, Michel H  
; APPLICANT: DU, Run-Pan  
; APPLICANT: EWASISHYN, Mary E  
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST  
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,189D  
; FILING DATE: 16-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jfb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

## US-08-838-189D-8

Query Match 100.0%; Score 1544; DB 2; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.1e-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
DB 1 MSKNKQORTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANEKVTLTAI IQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTOTTTILASTTP 120  
DB 61 FIASANEKVTLTAI IQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTOTTTILASTTP 120

QY 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQONKPNKPNNDHFEVFNFPVPCISCSNNP 180  
DB 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQONKPNKPNNDHFEVFNFPVPCISCSNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPTKKPTFKTKDLPQTKPKKEVPTTKPTTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKTTTKPTKKPTFKTKDLPQTKPKKEVPTTKPTTEPTINTTK 240

QY 241 TNITTTLLTNNNTGNPKLTSMETFHSTSEGNLSPSOVSTTSEHPSPQSPSPNTTRQ 298  
DB 241 TNITTTLLTNNNTGNPKLTSMETFHSTSEGNLSPSOVSTTSEHPSPQSPSPNTTRQ 298

## RESULT 3

US-08-852-344D-8  
; Sequence 8, Application US/08852344D  
; Patent No. 6017539  
; GENERAL INFORMATION:  
; APPLICANT: KLEIN, Michel H  
; APPLICANT: DU, Run-Pan  
; APPLICANT: EWASISHYN, Mary E  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION  
; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY  
; TITLE OF INVENTION: SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/852,344D  
; FILING DATE: 07-MAY-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/344,639  
; FILING DATE: 14-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-688 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

```
/
/ US-08-852-344D-8
/
/ TOPOLOGY: linear
/
/ Query Match 100.0%; Score 1544; DB 3; Length 298;
/ Best Local Similarity 100.0%; Pred. No. 2.1e-132;
/ Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY 1 MSKNKQORTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTSLITAI 60
/ Db 1 MSKNKQORTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTSLITAI 60
/
/ QY 61 FIASANKVLTALIIQDATSQIKNTPTTLTQDPOLGIFSLSNLSITSTTTILASTTP 120
/ Db 61 FIASANKVLTALIIQDATSQIKNTPTTLTQDPOLGIFSLSNLSITSTTTILASTTP 120
/
/ QY 121 GVKSNLOPTTVTKNTTTTQTSKPTTKQKQKPNKPNNDHFVFNVPSCSNNP 180
/ Db 121 GVKSNLOPTTVTKNTTTTQTSKPTTKQKQKPNKPNNDHFVFNVPSCSNNP 180
/
/ QY 181 TCWAIKCRIPNKKPGKTTTKPTKPTTKKDLKPQTKKEVPTTKPTTEPTINTTK 240
/ Db 181 TCWAIKCRIPNKKPGKTTTKPTKPTTKKDLKPQTKKEVPTTKPTTEPTINTTK 240
/
/ QY 241 TNITTTLLNTNNTGNPKLTSQMETFHSSTSEGNLSPSQSVSTTSEHPSQSPSPNTTRQ 298
/ Db 241 TNITTTLLNTNNTGNPKLTSQMETFHSSTSEGNLSPSQSVSTTSEHPSQSPSPNTTRQ 298
/
/ RESULT 4
/ US-08-344-639E-8
/ Sequence 8, Application US/08344639E
/ Patent No. 6033668
/ GENERAL INFORMATION:
/ APPLICANT: Klein, Michel H
/ APPLICANT: Du, Run-Pan
/ APPLICANT: Ewasysghyn, Mary E
/ TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
/ TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
/ TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
/ NUMBER OF SEQUENCES: 38
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sim & McBurney
/ STREET: 330 University Avenue, 6th Floor
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5G 1R7
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/344,639E
/ FILING DATE: 14-NOV-1994
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/001,554
/ FILING DATE: 06-JAN-1993
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9200117.1
/ FILING DATE: 06-JAN-1992
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stewart, Michael I
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-391 MIS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/ TELEFAX: (416) 595-1163
/ TELEX: 065-24567 SIMBAS
/ INFORMATION FOR SEQ ID NO: 8:
/
```

```
/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 298 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-344-639E-8
/
/ Query Match 100.0%; Score 1544; DB 3; Length 298;
/ Best Local Similarity 100.0%; Pred. No. 2.1e-132;
/ Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY 1 MSKNKQORTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTSLITAI 60
/ Db 1 MSKNKQORTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTSLITAI 60
/
/ QY 61 FIASANKVLTALIIQDATSQIKNTPTTLTQDPOLGIFSLSNLSITSTTTILASTTP 120
/ Db 61 FIASANKVLTALIIQDATSQIKNTPTTLTQDPOLGIFSLSNLSITSTTTILASTTP 120
/
/ QY 121 GVKSNLOPTTVTKNTTTTQTSKPTTKQKQKPNKPNNDHFVFNVPSCSNNP 180
/ Db 121 GVKSNLOPTTVTKNTTTTQTSKPTTKQKQKPNKPNNDHFVFNVPSCSNNP 180
/
/ QY 181 TCWAIKCRIPNKKPGKTTTKPTKPTTKKDLKPQTKKEVPTTKPTTEPTINTTK 240
/ Db 181 TCWAIKCRIPNKKPGKTTTKPTKPTTKKDLKPQTKKEVPTTKPTTEPTINTTK 240
/
/ QY 241 TNITTTLLNTNNTGNPKLTSQMETFHSSTSEGNLSPSQSVSTTSEHPSQSPSPNTTRQ 298
/ Db 241 TNITTTLLNTNNTGNPKLTSQMETFHSSTSEGNLSPSQSVSTTSEHPSQSPSPNTTRQ 298
/
/ RESULT 5
/ US-08-467-969A-8
/ Sequence 8, Application US/08467969A
/ Patent No. 6168786
/ GENERAL INFORMATION:
/ APPLICANT: Klein, Michel H
/ APPLICANT: Du, Run-Pan
/ APPLICANT: Ewasysghyn, Mary E
/ TITLE OF INVENTION: Chimeric Immunogens
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sim & McBurney
/ STREET: 330 University Avenue, 6th Floor
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5G 1R7
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/467,969A
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/001,554
/ FILING DATE: 06-JAN-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9200117.1
/ FILING DATE: 06-JAN-1992
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stewart, Michael I
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-475 MIS:bb
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/
```

```
;
; TELEFAX: (416) 595-1163
; TELEPHONE: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-467-961A-8

Query Match      100.0%; Score 1544; DB 3; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.1e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Db 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Qy 61 FIASANHKVLTITAIQDATSQIKNTTPTTYLTQDPQLGISFNSLSEITSTQTTILASTTP 120
Db 61 FIASANHKVLTITAIQDATSQIKNTTPTTYLTQDPQLGISFNSLSEITSTQTTILASTTP 120
Qy 121 GVKSNLQPTTVTKNTTTTQTPSKPTTKQKPNKPNNDFFHFEVNFVPCSCSNNP 180
Db 121 GVKSNLQPTTVTKNTTTTQTPSKPTTKQKPNKPNNDFFHFEVNFVPCSCSNNP 180
Qy 181 TCWAICRIKPNKPGKTKTKPTTKKPTTKKPKTKKQDKLPQTKPKVEPTTKTEPTINTTK 240
Db 181 TCWAICRIKPNKPGKTKTKPTTKKPTTKKPKTKKQDKLPQTKPKVEPTTKTEPTINTTK 240
Qy 241 TNITTTLLTNNNTGNPKLTQMETFHTSSEGNLSPSOVSTTSEHPSPSPPNNTTQ 298
Db 241 TNITTTLLTNNNTGNPKLTQMETFHTSSEGNLSPSOVSTTSEHPSPSPPNNTTQ 298

RESULT 6
US-08-467-961A-8
; Sequence 8, Application US/08467961A
; Patent No. 6171783
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysbyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,961A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
```

```
;
; REFERENCE/DOCKET NUMBER: 1038-476 MLS:bh
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELE: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-467-961A-8

Query Match      100.0%; Score 1544; DB 3; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.1e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Db 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Qy 61 FIASANHKVLTITAIQDATSQIKNTTPTTYLTQDPQLGISFNSLSEITSTQTTILASTTP 120
Db 61 FIASANHKVLTITAIQDATSQIKNTTPTTYLTQDPQLGISFNSLSEITSTQTTILASTTP 120
Qy 121 GVKSNLQPTTVTKNTTTTQTPSKPTTKQKPNKPNNDFFHFEVNFVPCSCSNNP 180
Db 121 GVKSNLQPTTVTKNTTTTQTPSKPTTKQKPNKPNNDFFHFEVNFVPCSCSNNP 180
Qy 181 TCWAICRIKPNKPGKTKTKPTTKKPTTKKPKTKKQDKLPQTKPKVEPTTKTEPTINTTK 240
Db 181 TCWAICRIKPNKPGKTKTKPTTKKPTTKKPKTKKQDKLPQTKPKVEPTTKTEPTINTTK 240
Qy 241 TNITTTLLTNNNTGNPKLTQMETFHTSSEGNLSPSOVSTTSEHPSPSPPNNTTQ 298
Db 241 TNITTTLLTNNNTGNPKLTQMETFHTSSEGNLSPSOVSTTSEHPSPSPPNNTTQ 298

RESULT 7
US-08-001-554A-8
; Sequence 8, Application US/08001554A
; Patent No. 6225091
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysbyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/001,554A
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-286
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELE: 065-24567 SIMBAS
```

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 298 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
US-08-001-554A-8

Query Match 100.0%; Score 1544; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.1e-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSVQITLSILAMISTLSIITAI 60  
DB 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSVQITLSILAMISTLSIITAI 60

QY 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPQLGISFSLSEITSTTILASTTP 120  
DB 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPQLGISFSLSEITSTTILASTTP 120

QY 121 GVKSNLOPTTVTKNTTITOTQPSKPTTKORQKPNKPNNDFFHFEVFNVPSCISNNP 180  
DB 121 GVKSNLOPTTVTKNTTITOTQPSKPTTKORQKPNKPNNDFFHFEVFNVPSCISNNP 180

QY 181 TCWAICKRIPNPKGKTTTKPTKPTKTKKDKLPQTTKPEVPTTKPTTEPTINTTK 240  
DB 181 TCWAICKRIPNPKGKTTTKPTKPTKTKKDKLPQTTKPEVPTTKPTTEPTINTTK 240

QY 241 TNIITTTLLTNNNTGNPKLTSQMETFHSSTSGNLSPSQSVSTSEHPSQSPSPNTTRQ 298  
DB 241 TNIITTTLLTNNNTGNPKLTSQMETFHSSTSGNLSPSQSVSTSEHPSQSPSPNTTRQ 298

RESULT 8  
5194595-19  
PATENT NO. 5194595  
APPLICANT: WATHEW, MICHAEL W.  
TITLE OF INVENTION: CHIMERIC GLYCOPROTEINS CONTAINING  
IMMUNOGENIC SEGMENT OF THE GLYCOPROTEINS OF HUMAN RESPIRATORY  
SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 19  
CURRENT APPLICATION DATA:  
FILING DATE: 31-OCT-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 137,387  
FILING DATE: 23-DEC-1987  
SEQ ID NO: 19:  
LENGTH: 681

Query Match 58.4%; Score 901; DB 6; Length 681;  
Best Local Similarity 91.0%; Pred. No. 1.6e-73;  
Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 96 QLGISFSLSEITSTTILASTTPGVKSNLOPTTVTKNTTITOTQPSKPTTKORQK 155  
DB 490 QLGISFSLSEITSTTILASTTPGVKSNLOPTTVTKNTTITOTQPSKPTTKORQK 155

QY 156 PNKPNNDFFHFEVFNVPSCISNNPCTWAI CKRIPNPKPGKTTTKPTKPTKTKKOL 215  
DB 550 PNKPNNDFFHFEVFNVPSCISNNPCTWAI CKRIPNPKPGKTTTKPTKPTKTKKOL 215

QY 216 KPQTTKPEVPTTKPTTEPTINTKNTTITLLTNNNTGNPKLTSQMETFHSSTSEGNLS 275  
DB 610 KPQTTKPEVPTTKPTTEPTINTKNTTITLLTNNNTGNPKLTSQMETFHSSTSEGNLS 275

QY 276 PSQVSTSE 284  
DB 670 PSQVNISSQ 678

RESULT 9  
PCT-US91-08177-13  
SEQUENCE 13, Application PC/TUS9108177  
GENERAL INFORMATION:  
APPLICANT: Samal, Siba K  
TITLE OF INVENTION: Bovine Respiratory Syncytial Virus Genes  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
STREET: 1201 New York Avenue N.W., suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08177  
FILING DATE: 19911104  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/608,937  
FILING DATE: 05-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Highest, David W  
REGISTRATION NUMBER: 30,265  
REFERENCE/DOCKET NUMBER: 20509-96711  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4854  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-08177-13

Query Match 19.1%; Score 295.5; DB 5; Length 263;  
Best Local Similarity 31.7%; Pred. No. 4.1e-19;  
Matches 86; Conservative 35; Mismatches 123; Indels 27; Gaps 7;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSVQITLSILAMISTLSIITAI 60  
DB 1 MSNTHHPKFKTLKRAWKASKYFIVGLSCLYKFNLSLVQATLSLMTSLTSLVITAI 60

QY 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPQLGISFSLSEITSTTILASTTP 120  
DB 61 YISVGNKAKPTSKPTTQQTQOPQNHIFLLPTEHNHKSHTSTSTTUSQPN1--DTTS 118

QY 121 GVK-----SNLOPTTVTKNT--TTTQTPSKPTTKORQKPNKPNNDFFHFEVFNVPSC 174  
DB 119 GTTYGHPINRTQNRKIKSQSTPLATKPLINPL----ESNPPENHQDNNQTLPHVPCS 174

QY 175 ICSNPTCWAIC-----KRPN-----KPGKTKTKPTKPTKTKKOLKPKQTKP 222  
DB 175 TCEGNPACPLCOIGLERAPSRAPTITLKAPKPKTKTKPTTKPTTKPTTKPTTK 234

QY 223 KEVPTTKPTTEPTINTKNTTITLLTNNNT 253  
DB 235 TATPQOGLISSP---EHQTNQSTTQISQHTS 262

RESULT 10  
US-08-793-792-12  
SEQUENCE 12, Application US/08793792  
PATENT NO. 6077511  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Antigenic peptides derived from the

;; TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis  
;; TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.  
;; NUMBER OF SEQUENCES: 13

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/793,792

;; FILING DATE:

;; CLASSIFICATION: 514

;; INFORMATION FOR SEQ ID NO: 12:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 37 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; HYPOTHETICAL: NO

US-08-793-792-12

Query Match 14.4%; Score 222; DB 3; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.3e-13;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 NKPNNDFHEVFNFPVPCISCSNNPTCWAICKRI 193

Db 1 NKPNNDFHEVFNFPVPCISCSNNPTCWAICKRI 37

RESULT 11

US-08-793-792-8

;; Sequence 8, Application US/08793792

;; Patent No. 6077511

;; GENERAL INFORMATION:

;; APPLICANT:

;; TITLE OF INVENTION: Antigenic peptides derived from the

;; TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis

;; TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.

;; NUMBER OF SEQUENCES: 13

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/793,792

;; FILING DATE:

;; CLASSIFICATION: 514

;; INFORMATION FOR SEQ ID NO: 8:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 32 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; HYPOTHETICAL: NO

US-08-793-792-8

Query Match 12.5%; Score 193; DB 3; Length 32;

Best Local Similarity 100.0%; Pred. No. 4.5e-11;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 KPNNDHEVFNFPVPCISCSNNPTCWAICKRI 189

Db 1 KPNNDHEVFNFPVPCISCSNNPTCWAICKRI 32

RESULT 12

US-08-928-361B-8

;; Sequence 8, Application US/08928361B

;; Patent No. 6071518

;; GENERAL INFORMATION:

;; APPLICANT: Petersen, Carolyn

;; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

;; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

;; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

;; TITLE OF INVENTION: SPECIES INFECTIONS

;; NUMBER OF SEQUENCES: 30

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: PETERS, VERNY, JONES & BIKSA

;; STREET: 385 Sherman Avenue, Suite 6

;; CITY: Palo Alto

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 94306-1840

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/928,361B

;; FILING DATE: 12-SEP-1997

;; CLASSIFICATION:

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 60/026,062

;; FILING DATE: 13-SEP-1996

;; ATTORNEY/AGENT INFORMATION:

;; NAME: VERNY, Hana

;; REGISTRATION NUMBER: 30,518

;; REFERENCE/DOCKET NUMBER: 480.76-1(HV)

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 650-324-1677

;; TELEFAX: 650-324-1678

;; INFORMATION FOR SEQ ID NO: 8:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 216 amino acids

;; TYPE: amino acid

;; STRANDEDNESS:

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

US-08-928-361B-8

Query Match 12.2%; Score 188; DB 3; Length 216;

Best Local Similarity 26.4%; Pred. No. 1.8e-09;

Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;

QY 70 TLTAIIQDATSQIKNTPTPLTQDPOLGISFNSLSEITQSITLILASTTPGVKSNLOPT 129

Db 12 TTTTITT 71

QY 130 TVTKNTTTTQTPSKPTTKORONKPNKPNNDHFHFEVFNFPVPCISCSNNPTCWAICKRI 189

Db 72 TTTTITT 111

QY 190 PNKKPGKTTTKKPKTKKDLKPKQTKPKVPTTKPTKEPTINTKNTLT 249

Db 112 TTKKPTT 171

QY 250 NNTTGNPKLTSMETHTSSEGNLSPQSVSTTSEHPSQSPSPNTT 296

Db 172 TTTTITTTTTTTTTTKKPTTT-----TTTTTTTTTKKPTTTATTTT 214

RESULT 13

US-09-588-995A-8

;; Sequence 8, Application US/09588995A

;; Patent No. 6514697

;; GENERAL INFORMATION:

;; APPLICANT: PETERSEN, CAROLYN

;; APPLICANT: BARNES, DEBRA A.

;; APPLICANT: NELSON, RICHARD C.

;; APPLICANT: GUT, JIRI

;; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND





Search completed: October 29, 2003, 17:42:22  
Job time : 17.3057 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:11:05 ; Search time 68.034 Seconds  
(without alignments)  
695.249 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDQRTAKTLEKTDTL.....VSTTSEHPSPSPNPTTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq 19Jun03.\*

|     |  |
|-----|--|
| 1:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.* |
| 2:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.* |
| 3:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.* |
| 4:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.* |
| 5:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.* |
| 6:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.* |
| 7:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.* |
| 8:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.* |
| 9:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.* |
| 10: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.* |
| 11: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.* |
| 12: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.* |
| 13: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.* |
| 14: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.* |
| 15: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.* |
| 16: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.* |
| 17: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.* |
| 18: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.* |
| 19: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.* |
| 20: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.* |
| 21: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.* |
| 22: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.* |
| 23: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.* |
| 24: | /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 1544  | 100.0       | 298    | 14 | Respiratory syncyt |
| 2          | 1544  | 100.0       | 298    | 20 | Membrane bound G p |
| 3          | 1418  | 91.8        | 298    | 8  | Sequence of human  |
| 4          | 1418  | 91.8        | 298    | 13 | HSRV glycoprotein  |
| 5          | 1418  | 91.8        | 298    | 19 | HSRV glycoprotein  |
| 6          | 1418  | 91.8        | 298    | 23 | Respiratory syncyt |
| 7          | 1229  | 79.6        | 232    | 20 | Secreted G protein |
| 8          | 901   | 58.4        | 681    | 10 | Chimeric human res |
| 9          | 771   | 49.9        | 299    | 22 | Amino acid sequenc |

|    |       |      |     |    |          |                    |
|----|-------|------|-----|----|----------|--------------------|
| 10 | 580   | 37.6 | 361 | 24 | ABP97862 | RSV G protein pept |
| 11 | 572.5 | 37.1 | 452 | 22 | ABP67771 | Amino acid sequenc |
| 12 | 564.5 | 36.6 | 291 | 24 | ABP97861 | RSV G protein pept |
| 13 | 564.5 | 36.6 | 548 | 24 | ABP97863 | RSV G protein pept |
| 14 | 562   | 36.4 | 349 | 17 | AA995660 | Streptococcal prot |
| 15 | 562   | 36.4 | 349 | 22 | AA995660 | Amino acid sequenc |
| 16 | 558   | 36.1 | 101 | 16 | AA988253 | RSV subgroup A wil |
| 17 | 558   | 36.1 | 101 | 17 | AA995610 | RSV subgp. A prote |
| 18 | 558   | 36.1 | 101 | 17 | AA995616 | RSV sub-group A wi |
| 19 | 558   | 36.1 | 101 | 17 | AA997050 | Respiratory Syncyt |
| 20 | 558   | 36.1 | 101 | 21 | AA944078 | RSV G protein anti |
| 21 | 558   | 36.1 | 101 | 21 | AA944078 | A G2Na peptide der |
| 22 | 558   | 36.1 | 101 | 22 | AA967741 | Amino acid sequenc |
| 23 | 558   | 36.1 | 101 | 22 | AA984123 | Amino acid sequenc |
| 24 | 558   | 36.1 | 101 | 22 | AA984123 | Amino acid sequenc |
| 25 | 558   | 36.1 | 101 | 22 | AA967775 | Human G protein, G |
| 26 | 558   | 36.1 | 101 | 23 | AAO22577 | Human G protein, G |
| 27 | 552.5 | 35.8 | 356 | 17 | AA995661 | Immunogenic carrie |
| 28 | 538   | 34.8 | 101 | 16 | AA988255 | RSV subgroup A mod |
| 29 | 538   | 34.8 | 101 | 17 | AA995612 | RSV subgp. A prote |
| 30 | 538   | 34.8 | 101 | 17 | AA995618 | RSV sub-group A mu |
| 31 | 538   | 34.8 | 101 | 17 | AA997052 | Respiratory Syncyt |
| 32 | 538   | 34.8 | 101 | 20 | AA944080 | RSV G protein anti |
| 33 | 538   | 34.8 | 101 | 20 | AA997311 | Peptide which indu |
| 34 | 538   | 34.8 | 101 | 22 | AA984125 | Amino acid sequenc |
| 35 | 538   | 34.8 | 101 | 23 | AAO22581 | Human G protein, G |
| 36 | 538   | 34.8 | 101 | 23 | AAO22586 | Human G protein, G |
| 37 | 534   | 34.6 | 101 | 20 | AA997310 | Peptide which indu |
| 38 | 534   | 34.6 | 101 | 23 | AAO22585 | Human G protein, G |
| 39 | 514   | 33.3 | 101 | 20 | AA997312 | Peptide which indu |
| 40 | 514   | 33.3 | 101 | 23 | AAO22587 | Human G protein, G |
| 41 | 506   | 32.8 | 101 | 17 | AA995614 | RSV subgp. A prote |
| 42 | 506   | 32.8 | 101 | 17 | AA997063 | Respiratory Syncyt |
| 43 | 506   | 32.8 | 101 | 20 | AA944090 | RSV G protein anti |
| 44 | 506   | 32.8 | 101 | 22 | AA984135 | Amino acid sequenc |
| 45 | 506   | 32.8 | 101 | 23 | AAO22584 | Human G protein, G |

#### ALIGNMENTS

#### RESULT 1

AA939286  
ID AAR39286 standard; Protein; 298 AA.

XX AAR39286;

XX 25-MAR-2003 (updated)

DT 13-JAN-1994 (first entry)

XX Respiratory syncytial virus (RSV) G protein.

DE Respiratory syncytial virus (RSV) G protein.

XX PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.

OS Respiratory syncytial virus.

XX WO9314207-A1.

XX 22-JUL-1993.

XX 05-JAN-1993; 93WO-CA00001.

XX 06-JAN-1992; 92GB-0000117.

XX (CONN-) CONNAUGHT LAB LTD.

XX Ewasysbyn ME, Klein MH;

XX WPI; 1993-243222/30.

XX N-PSDB; AAQ45686.

PT Multimeric hybrid genes and their chimeric proteins - are

PT vaccines against multiple pathogenic infections e.g.

PT para-influenza virus and respiratory syncytial virus  
 XX Claim 11; Figure 7A-7D; 80pp; English.

XX A novel multimeric hybrid gene is used as a vaccine. The gene  
 CC consists of two gene sequences which are linked and encode antigenic  
 CC regions, these two sequences being derived from two different  
 CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus  
 CC (RSV)). The gene sequences that are particularly used are those  
 CC which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
 CC G proteins (AAQ45685, AAQ45686).  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 298 AA;

Query Match 100.0%; Score 1544; DB 14; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-113;  
 Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTDTLNHLLFISSGLYKLNKLSVAQITLSILAMIIISTLIITAI 60  
 DB 1 MSKNKQORTAKTLEKWTDTLNHLLFISSGLYKLNKLSVAQITLSILAMIIISTLIITAI 60  
 QY 61 FIASANKHVLTITAIQDATSQIKNTTPTLTQDPQLGIFSLSNLSITTSOTTTILASTTP 120  
 DB 61 FIASANKHVLTITAIQDATSQIKNTTPTLTQDPQLGIFSLSNLSITTSOTTTILASTTP 120  
 QY 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEFVFNVPVCSICSNNP 180  
 DB 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEFVFNVPVCSICSNNP 180  
 QY 181 TCWAICKRIPNKKPGKKTTKTKPTTKTKKDLKPQTKPKVEPTTKPTBPTINTTK 240  
 DB 181 TCWAICKRIPNKKPGKKTTKTKPTTKTKKDLKPQTKPKVEPTTKPTBPTINTTK 240  
 QY 241 TNITTTLLTNNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSPQSSPPNTTRQ 298  
 DB 241 TNITTTLLTNNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSPQSSPPNTTRQ 298

#### RESULT 2

AAW96313  
 ID AAW96313 standard; Protein; 298 AA.  
 XX  
 AC AAW96313;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Membrane bound G protein of respiratory syncytial virus.  
 DE  
 DE G protein; respiratory syncytial virus; RSV; recombinant vector;  
 KW vaccine; immune response; immunogenicity; tPA; antibody;  
 KW tissue plasminogen activator.  
 XX  
 XX Respiratory syncytial virus.

WO9904010-A1.

28-JAN-1999.

16-JUL-1998; 98WO-CA00697.

18-JUL-1997; 97US-0896442.

(CONN-) CONNAUGHT LAB LTD.

Klein MH, Li X, Sambhara S;

WPI; 1999-132254/11.

DR N-PSDB; AAX08421.

XX Immunogenic composition for generating antibodies against  
 PT respiratory syncytial virus - comprises non-replicating vector

PT containing the protein G sequence, useful in protective vaccines and  
 PT to raise antibodies for diagnosis  
 XX Claim 4; Fig 2; 67pp; English.

XX The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods.

SQ Sequence 298 AA;

Query Match 100.0%; Score 1544; DB 20; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-113;  
 Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTDTLNHLLFISSGLYKLNKLSVAQITLSILAMIIISTLIITAI 60  
 DB 1 MSKNKQORTAKTLEKWTDTLNHLLFISSGLYKLNKLSVAQITLSILAMIIISTLIITAI 60  
 QY 61 FIASANKHVLTITAIQDATSQIKNTTPTLTQDPQLGIFSLSNLSITTSOTTTILASTTP 120  
 DB 61 FIASANKHVLTITAIQDATSQIKNTTPTLTQDPQLGIFSLSNLSITTSOTTTILASTTP 120  
 QY 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEFVFNVPVCSICSNNP 180  
 DB 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEFVFNVPVCSICSNNP 180  
 QY 181 TCWAICKRIPNKKPGKKTTKTKPTTKTKKDLKPQTKPKVEPTTKPTBPTINTTK 240  
 DB 181 TCWAICKRIPNKKPGKKTTKTKPTTKTKKDLKPQTKPKVEPTTKPTBPTINTTK 240  
 QY 241 TNITTTLLTNNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSPQSSPPNTTRQ 298  
 DB 241 TNITTTLLTNNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSPQSSPPNTTRQ 298

#### RESULT 3

AAP70845  
 ID AAP70845 standard; protein; 298 AA.  
 XX  
 AC AAP70845;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 05-APR-1991 (first entry)  
 XX  
 XX Sequence of human respiratory syncytial virus (HRSV) A2 strain  
 DE G protein.  
 DE  
 XX Vaccine.  
 XX  
 XX Human respiratory syncytial virus (HRSV).  
 XX  
 XX WO8704185-A.

16-JUL-1987.

23-DEC-1986; 86WO-US02756.

14-JAN-1986; 86US-0818740.

(UNIC-) UNIV NORTH CAROLINA.

PA (WERTZ) WERTZ G W.

```

DR WPI; 1987-206300/29.
DR N-PSDB; AAN70784.
XX Vaccines for human respiratory virus - comprising proteins or
PT fragment encoded by a DNA sequence coding for human respiratory
PT syncytial virus proteins.
XX Disclosure; Chart 13; 57pp; English.
XX A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 298 AA;
SQ Query Match 91.8%; Score 1418; DB 8; Length 298;
Best Local Similarity 93.3%; Pred. No. 2e-103;
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNKQORTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANKHVTLTALIQDATSQIKNTPTTYLTQDPOLGISFNSLSEITSTTILASTTP 120
DB 61 FIASANKHVTPTTALIQDATSQIKNTPTTYLTQDPOLGISFNSLSEITSTTILASTTP 120

QY 121 GVKSNLQPTTVKTKNTTTTQPSKPTTKORQNKPNKPNNDHFEVFNVPSCISNNP 180
DB 121 GVKSTLQSTTVKTKNTTTTQPSKPTTKORQNKPNKPNNDHFEVFNVPSCISNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPKPTTKKDLKPQTTKPKKEVPTTKPTTEPTINTTK 240
DB 181 TCWAICKRIPNKKPGKTTTKPKPTTKKDLKPQTTKPKKEVPTTKPTTEPTINTTK 240

QY 241 TNIITTTLLTNTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298
DB 241 TNIITTTLLTNTGNPELTSQMETFHTSSEGNLSPSQVSTTSEYPSQSPSPNTPRQ 298

RESULT 4
AAR25302
ID AAR25302 standard; Protein; 298 AA.
XX AC AAR25302;
XX AC AAR25302;
DT 25-MAR-2003 (updated)
DT 03-MAR-1993 (first entry)
XX DE HRSV glycoprotein G (gpG).
XX KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
KW major capsid protein; N.
XX OS Human respiratory syncytial virus strain A2.
XX PN US5149650-A.
XX PD 22-SEP-1992.
XX PF 13-JUL-1988; 88US-0218737.
XX PR 14-JAN-1986; 86US-0818740.
XX PR 13-JUL-1988; 88US-0218737.
XX PA (UYN-C-) UNIV NORTH CAROLINA.
XX PI Collins PL, Wertz GW;
XX WPI; 1992-340247/41.

DR N-PSDB; AAQ29623.
XX Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
PT fragments
XX Disclosure; Page 18; 21pp; English.
XX The sequences of mRNA encoding HRSV structural proteins are given in
CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid
CC protein N. The sequences and encoded proteins are useful for
CC preparing vaccines against HRSV. The vaccines can be used to confer
CC immunity against respiratory tract infections on human subjects.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX Sequence 298 AA;
SQ Query Match 91.8%; Score 1418; DB 13; Length 298;
Best Local Similarity 93.3%; Pred. No. 2e-103;
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNKQORTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANKHVTLTALIQDATSQIKNTPTTYLTQDPOLGISFNSLSEITSTTILASTTP 120
DB 61 FIASANKHVTPTTALIQDATSQIKNTPTTYLTQDPOLGISFNSLSEITSTTILASTTP 120

QY 121 GVKSNLQPTTVKTKNTTTTQPSKPTTKORQNKPNKPNNDHFEVFNVPSCISNNP 180
DB 121 GVKSTLQSTTVKTKNTTTTQPSKPTTKORQNKPNKPNNDHFEVFNVPSCISNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPKPTTKKDLKPQTTKPKKEVPTTKPTTEPTINTTK 240
DB 181 TCWAICKRIPNKKPGKTTTKPKPTTKKDLKPQTTKPKKEVPTTKPTTEPTINTTK 240

QY 241 TNIITTTLLTNTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298
DB 241 TNIITTTLLTNTGNPELTSQMETFHTSSEGNLSPSQVSTTSEYPSQSPSPNTPRQ 298

RESULT 5
AAW47605
ID AAW47605 standard; Protein; 298 AA.
XX AC AAW47605;
XX AC AAW47605;
DT 11-JUN-1998 (first entry)
XX DE HRSV glycoprotein G.
XX KW HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine.
XX OS Human respiratory syncytial virus.
XX PN US5716823-A.
XX PD 10-FEB-1998.
XX PF 12-MAY-1997; 97US-0854783.
XX PR 13-JUL-1988; 88US-0218737.
XX PR 14-JAN-1986; 86US-0818740.
XX PR 23-DEC-1986; 86WO-US02756.
XX PR 11-JUN-1992; 92US-0897171.
XX PR 12-MAY-1997; 97US-0854783.
XX PA (PHAA ) PHARMACIA & UPJOHN CO.
XX PI Collins PL, Wertz GW;
XX WPI; 1998-144802/13.

```

DR N-PSDB; AAV18736.  
 XX Production of human respiratory syncytial virus glyco-protein F or G  
 PT - by culturing eukaryotic host cells transfected with corresponding  
 PT DNA  
 XX  
 XX Example 1; Columns 27-28; 17pp; English.  
 XX  
 CC The present sequence was used in the development of a novel method  
 CC for the production of human respiratory syncytial virus (HRSV)  
 CC glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises  
 CC culturing eukaryotic host cells transfected with an isolated DNA  
 CC sequence encoding HRSV gpF or gpG. The gp can be used to prepare  
 CC vaccines against HRSV.  
 XX  
 XX Sequence 298 AA;  
 SQ  
 Query Match 91.8%; Score 1418; DB 19; Length 298;  
 Best Local Similarity 93.3%; Pred. No. 2e-103;  
 Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
 DB 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
 QY 61 FIASANHKVTITTAIIQDATSQIKNTTPTVLTQDQPGISFNSLSEITTSQTTILASTTP 120  
 DB 61 FIASANHKVTITTAIIQDATSQIKNTTPTVLTQDQPGISFNSLSEITTSQTTILASTTP 120  
 QY 121 GVKSNLOPTTVTKNTTQTQPSKPTTKORONKPNKPNNDHFEVFNVPSCSNNP 180  
 DB 121 GVKSTLQSTTVTKNTTQTQPSKPTTKORONKPNKPNNDHFEVFNVPSCSNNP 180  
 QY 181 TCWAICRIPNKKPGKTTTKPTKPTKPTTKKOLKPKQTTKPKVPTTKTEPTINTTK 240  
 DB 181 TCWAICRIPNKKPGKTTTKPTKPTKPTTKKOLKPKQTTKPKVPTTKTEPTINTTK 240  
 QY 241 TNITTTLLTNTTGNPKLTQMETFHSSTSEGNLSPSQVSTTSHPSPSPSPNTTRQ 298  
 DB 241 TNITTTLLTNTTGNPKLTQMETFHSSTSEGNLSPSQVSTTSHPSPSPSPNTTRQ 298  
 RESULT 6  
 AAU74676  
 ID AAU74676 standard; Protein; 298 AA.  
 AC AAU74676;  
 DT 09-APR-2002 (first entry)  
 DE Respiratory syncytial virus G protein.  
 KW RSV; G protein; heavily glycosylated protein; antianaemic; antiviral;  
 KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;  
 KW antiviral chemotherapeutic compound; humoral response;  
 KW cellular immune response; hPIV; paediatric respiratory disease;  
 KW globin gene transfer; sickle cell disease; beta-thalassaemia;  
 KW human immunodeficiency virus infection; HIV.  
 XX  
 OS Human respiratory syncytial virus.  
 XX  
 PN WO200192548-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 22-MAY-2001; 2001WO-US16610.  
 XX  
 PR 01-JUN-2000; 2000US-208701P.  
 XX  
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Portner A, Takimoto T;  
 XX

DR WPI; 2002-130534/17.  
 XX N-PSDB; AAS21045.  
 PT Recombinant Sendai virus useful in vaccines to protect infection by  
 PT paramyxoviruses, comprises exogenous nucleic acid encoding  
 PT paramyxovirus protein or its antigenic fragment -  
 XX  
 XX Disclosure; Page 48; 57pp; English.  
 XX  
 CC The invention relates to a recombinant Sendai virus comprising an  
 CC exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its  
 CC antigenic fragment. The virus may be administered in combination  
 CC with an antiviral chemotherapeutic compound. Two or more viruses  
 CC expressing different PMV proteins may be co-administered. Compositions  
 CC comprising the virus are useful for eliciting a humoral and/or  
 CC cellular immune response to a PMV in a mammal, particularly a human.  
 CC Further a recombinant Sendai virus comprising an exogenous nucleic acid  
 CC encoding a second PMV protein is also administered and priming and/or  
 CC boosting humoral or cellular immune response comprises administering  
 CC one or more of a recombinant or isolated PMV protein or its antigenic  
 CC fragment, a DNA vaccine encoding the same, and a non-Sendai viral  
 CC vector encoding a PMV protein. The recombinant virus is useful as an  
 CC effective vaccine against hPIV or RSV (the major causes of paediatric  
 CC respiratory disease) and also to express any gene of  
 CC interest in target cells, providing a positive medical impact on  
 CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)  
 CC into stem cells effects a cure for sickle cell disease or beta-  
 CC thalassaemia. The recombinant virus may also prove effective in  
 CC conferring immunity to human immunodeficiency virus (HIV) infection.  
 CC The Sendai virus replicates at level that is high enough to  
 CC induce sufficient immunity, but does not cause any harm to human  
 CC recipient. The present sequence represents a respiratory syncytial  
 CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein  
 CC suitable for expression by the recombinant virus of the invention.  
 XX  
 XX Sequence 298 AA;  
 SQ  
 Query Match 91.8%; Score 1418; DB 23; Length 298;  
 Best Local Similarity 93.3%; Pred. No. 2e-103;  
 Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
 DB 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
 QY 61 FIASANHKVTITTAIIQDATSQIKNTTPTVLTQDQPGISFNSLSEITTSQTTILASTTP 120  
 DB 61 FIASANHKVTITTAIIQDATSQIKNTTPTVLTQDQPGISFNSLSEITTSQTTILASTTP 120  
 QY 121 GVKSNLOPTTVTKNTTQTQPSKPTTKORONKPNKPNNDHFEVFNVPSCSNNP 180  
 DB 121 GVKSTLQSTTVTKNTTQTQPSKPTTKORONKPNKPNNDHFEVFNVPSCSNNP 180  
 QY 181 TCWAICRIPNKKPGKTTTKPTKPTKPTTKKOLKPKQTTKPKVPTTKTEPTINTTK 240  
 DB 181 TCWAICRIPNKKPGKTTTKPTKPTKPTTKKOLKPKQTTKPKVPTTKTEPTINTTK 240  
 QY 241 TNITTTLLTNTTGNPKLTQMETFHSSTSEGNLSPSQVSTTSHPSPSPSPNTTRQ 298  
 DB 241 TNITTTLLTNTTGNPKLTQMETFHSSTSEGNLSPSQVSTTSHPSPSPSPNTTRQ 298  
 RESULT 7  
 AAU96314  
 ID AAU96314 standard; Protein; 232 AA.  
 XX  
 AC AAU96314;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Secreted G protein of respiratory syncytial virus.  
 XX  
 KW G protein; respiratory syncytial virus; RSV; recombinant vector;



```

PD 26-APR-2001.
XX
PF 23-OCT-2000; 2000WO-GB04084.
XX
PR 21-OCT-1999; 99GB-0024990.
XX
PA (GLAX ) GLAXO GROUP LTD.
XX
XX Malhotra R, Bird M;
XX
DR WPI; 2001-316238/33.
XX
PT Treatment of respiratory syncytial virus infection (RSV) using Annexin
XX II or L-selectin derivatives -
XX
PS Disclosure; Page 71-73; 74pp; English.
XX
XX The specification describes a product which antagonizes binding of
XX respiratory syncytial virus (RSV) G-protein to annexin II or
XX L-selectin, or causes a decrease in cell surface levels of annexin
XX II or L-selectin. Such products are used in the manufacture of a
XX medicament for use in prevention or treatment of RSV infection.
XX The present sequence represents a RSV G-protein.
XX
SQ Sequence 299 AA;
Query Match 49.9%; Score 771; DB 22; Length 299;
Best Local Similarity 54.7%; Pred. No. 1.3e-52;
Matches 162; Conservative 36; Mismatches 96; Indels 2; Gaps 2;
QY 1 MSKKNQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLIITAI 60
DB 1 MSKKNQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSLAMIISTSLIITAI 60
QY 61 FIASANKHVTLTALIQDATSQIKNTTPTLYTQDPQLGISFNLSEITTSQTTLASTTP 120
DB 61 FIASANKHVTLTAVTVQIKHTEKNITLYTQVPPERVSSKQPTTTSPIHTNSATSP 120
QY 121 GVKSNLQPTTKTKNTTQTPSPKPTTKQKQKPNKPNNDHFVEFVPCISCSNPP 180
DB 121 NTKSETHHTAQTKGRITTSQTNKPSTKPLKNPKPKKDDYHFEFVPCISGNNQ 180
QY 181 TCWAICKRIPNKKCKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 239
DB 181 LCKSICKTIPNKKPKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
QY 240 KNTITTTLLTNNNTGNPKLTSMQTFHSTSEGNLSPSQVSTTSHSPQSPPT 295
DB 241 ERDTSQSTVLDTTLEHTIQOQSLHSTTPENTPNSQTPTASE-PSTSNSTQNT 295

RESULT 10
ABP97862
ID ABP97862 standard; protein; 361 AA.
XX
AC ABP97862;
XX
XX 03-JUN-2003 (first entry)
XX
DE RSV G protein peptide G2Na and diphtheria anatoxin peptide DTb.
XX
KW Diphtheria anatoxin; immune response; antigen; carrier; G protein;
KW hypersensitivity response; vaccine; infection; RSV; cancer.
XX
OS Synthetic.
OS Respiratory syncytial virus.
OS Corynebacterium diphtheriae.
XX
FN FR2827606-A1.
XX
PD 24-JAN-2003.
XX
PF 20-JUL-2001; 2001FR-0009733.

```

```

XX 20-JUL-2001; 2001FR-0009733.
XX
XX (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
XX Corvaia N, Nguyen NT, Beck A;
XX
XX WPI; 2003-241969/25.
XX
XX New peptide derived from diphtheria anatoxin, useful as carrier in
XX vaccines, lacks at least one Cys residue, also related nucleic acids
XX
XX Disclosure; Page 27-28; 42pp; French.
XX
XX The present sequence represents a fusion protein comprising a peptide
XX derived from the G protein of Respiratory syncytial virus (RSV) linked
XX to a peptide derived from a diphtheria anatoxin. The diphtheria anatoxin
XX peptide generates an immune response against any antigen coupled to it,
XX and has a negative, immediate hypersensitivity response. Diphtheria
XX anatoxin derived peptides, when modified to lack at least one cysteine
XX residue, are useful as carrier peptides. Deletion of Cys residues in
XX anatoxin peptides reduces formation of unwanted disulfide bridges. The
XX peptides are used as a carrier for vaccines, particularly those for
XX prevention or treatment of viral, bacterial, parasitic or fungal
XX infections, or cancers and to generate, or increase, an immune response
XX against infectious agents or tumour cells.
XX
SQ Sequence 361 AA;
Query Match 37.6%; Score 580; DB 24; Length 361;
Best Local Similarity 70.8%; Pred. No. 1.6e-37;
Matches 114; Conservative 5; Mismatches 20; Indels 22; Gaps 3;
QY 130 TVTKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFVPCISCSNNTCWAICKRI 189
DB 4 TVTKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFVPCISCSNNTCWAICKRI 63
QY 190 PNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 241
DB 64 PNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 120
QY 242 NIIT-----TLLTNTTGNPKLTSMQTFHSTSE 271
DB 121 KIESLKEHGPIKNKMSSEPNKTVSEKAKQYLEEFHOTALE 161

RESULT 11
AAB67771
ID AAB67771 standard; Protein; 452 AA.
XX
AC AAB67771;
XX
XX 11-JUN-2001 (first entry)
XX
DE Amino acid sequence of a fusion protein of P40 and RSV antigen.
XX
KW Outer membrane protein A; OmpA; P40; enterobacteria; nasal composition;
KW RSV; RSV infection; lung; respiratory tract; vaccine.
XX
OS Synthetic.
OS Klebsiella pneumoniae.
OS Respiratory syncytial virus.
XX
FN WO200121203-A1.
XX
XX 29-MAR-2001.
XX
XX 22-SEP-2000; 2000WO-FR02626.
XX
XX 23-SEP-1999; 99FR-0011888.
XX
XX (FABR ) FABRE MEDICAMENT SA PIERRE.
XX

```





CC to a peptide derived from a diphtheria anatoxin. The diphtheria anatoxin  
 CC peptide generates an immune response against any antigen coupled to it,  
 CC and has a negative, immediate hypersensitivity response. Diphtheria  
 CC anatoxin derived peptides, when modified to lack at least one cysteine  
 CC residue, are useful as carrier peptides. Deletion of Cys residues in  
 CC anatoxin peptides reduces formation of unwanted disulfide bridges. The  
 CC peptides are used as a carrier for vaccines, particularly those for  
 CC prevention or treatment of viral, bacterial, parasitic or fungal  
 CC infections, or cancers and to generate, or increase, an immune response  
 CC against infectious agents or tumour cells.

XX SQ Sequence 548 AA;

Query Match 36.6%; Score 564.5; DB 24; Length 548;  
 Best Local Similarity 67.1%; Pred. No. 4.4e-36;  
 Matches 110; Conservative 10; Mismatches 21; Indels 23; Gaps 3;  
 QY 130 TVTKNTTTTQTPSKPTTKQKQKPPNPNDFEVEFVPCISCSNNPTCWAICKRI 189  
 Db 4 TVTKNTTTTQTPSKPTTKQKQKPPNPNDFEVEFVPCISCSNNPTCWAICKRI 63  
 QY 190 PNKKPGKTTTKPTKPTKTKKDKLKPQTTKPEVPTTKPTE--EPTINTTKNTLTTL 247  
 Db 64 PNKKPGKTTTKPTKPTKPTKTKKDKLKPQTTKPEVPTTKPTE--EPTINTTKNTLTTL 247  
 QY 248 LTNNTTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQPS 291  
 Db 120 -----MENF---SSYHGTKFGYVDSIQGIQKPKS 146

RESULT 14  
 AAR95660  
 ID AAR95660 standard; Protein; 349 AA.

XX AC AAR95660;  
 XX DT 21-JAN-1997 (first entry)  
 XX DE Streptococcal protein G fragment BB fused to RSV haptan G2A.  
 KW Streptococcus; protein G; carrier protein; immunogenicity; increase;  
 KW enhance; vaccine; anti-viral; human serum albumin; binding;  
 KW immunostimulation; respiratory syncytial virus; fusion protein.  
 XX OS Chimeric Streptococcus sp.  
 OS Chimeric Respiratory Syncytial Virus.  
 XX FH Key Location/Qualifiers  
 FT Peptide 1..23  
 FT /label= signal\_peptide  
 FT Protein 24..242  
 FT /label= BB  
 FT Protein 247..347  
 FT /label= G2A  
 FT /note= "Residues 130-230 of RSV"  
 XX WO9614416-A1.  
 XX PD 17-MAY-1996.  
 XX PF 07-NOV-1995; 95WO-FR01466.  
 XX PR 07-NOV-1994; 94FR-0013310.  
 XX PA (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX Andreoni C, Binz H, Nguyen Ngoc T, Nygren PA, Stahl S;  
 PI Uhlen M;  
 DR WPI; 1996-251766/25.  
 DR N-PSDB; AAT31647.  
 XX Enhancing immunogenicity by coupling immunogen to serum

PT albumin-binding protein - useful for preparing improved vaccines,  
 PT e.g. against Respiratory Syncytial Virus  
 XX PS Disclosure; Page 81-83; 102pp; French.  
 XX CC The present sequence comprises a 219 amino acid fragment of  
 CC Streptococcal protein G, designated BB, fused to an immunogen,  
 CC designated G2A, derived from amino acids 130-230 of protein G from  
 CC Respiratory Syncytial Virus sub-group A or B. The BB fragment is  
 CC able to bind to human serum albumin and thereby enhance immunogenicity  
 CC of any antigen, hapten or immunogen that is covalently coupled to it.  
 CC In this specific example, the BB fragment was found to induce T helper  
 CC memory cells, leading to production of anti-G2A antibodies by stimulated  
 CC B cells.

XX SQ Sequence 349 AA;

Query Match 36.4%; Score 562; DB 17; Length 349;  
 Best Local Similarity 78.1%; Pred. No. 4e-36;  
 Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;  
 QY 103 NLSEITSQTT--ILASTT-----PGVKSNIQPTTVTKNTTTTQTPSKPTTKORQNK 155  
 Db 216 NGKTLKGTTTEAVDAATARSFNPFLENSM---TVTKNTTTTQTPSKPTTKORQNK 272  
 QY 156 PNKNPNDFHFEVFPVPCISCSNNPTCWAICKRIPNKKPKGKTTTKPKTEFKTKKOL 215  
 Db 273 PNKNPNDFHFEVFPVPCISCSNNPTCWAICKRIPNKKPKGKTTTKPKTEFKTKKOL 215  
 QY 216 KPQTTKPEVPTTKPTE 232  
 Db 333 KPQTTKPEVPTTKPTE 349

RESULT 15  
 AAB68028  
 ID AAB68028 standard; Protein; 349 AA.  
 XX AC AAB68028;  
 XX DT 29-JUN-2001 (first entry)  
 XX DE Amino acid sequence of fusion protein comprising 2 G protein fragments.  
 XX KW Aliphatic ammonium salt; immunogen; antigen; syncytial virus infection;  
 KW G protein; fusion protein.  
 XX OS Synthetic.  
 OS Streptococcus sp.  
 OS Respiratory syncytial virus.  
 XX FH Key Location/Qualifiers  
 FT Protein 120..230  
 FT /note= "G protein fragment of VRS"  
 XX FR2798292-A1.  
 XX PD 16-MAR-2001.  
 XX PF 09-SEP-1999; 99FR-0011284.  
 XX PR 09-SEP-1999; 99FR-0011284.  
 XX PA (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX Beck A, Klinguer C, Nguyen TN;  
 PI WPI; 2001-267782/28.  
 DR N-PSDB; AAF84711.  
 XX Use of quaternary aliphatic ammonium salt and immunogen or antigen to  
 PT combat respiratory syncytial virus infections



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 21:18:11 ; Search time 1944.71 Seconds  
(without alignments)  
8935.871 Million cell updates/sec

Title: US-09-462-816-3  
Perfect score: 715  
Sequence: 1 cacaagaagtcactaacac.....gtagttattataaaaaaaaaa 715

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| C 1        | 120   | 16.8        | 895    | 29 | CNS0071A            |
| C 2        | 113   | 15.8        | 712    | 13 | AL066286 Drosophila |
| C 3        | 111.4 | 15.6        | 1180   | 13 | BX416727            |
| C 4        | 104.6 | 14.6        | 946    | 29 | BX436369 BX436369   |
|            |       |             |        |    | AG127412 Pan trogl  |

|      |       |      |      |    |          |
|------|-------|------|------|----|----------|
| C 5  | 103.2 | 14.4 | 922  | 29 | CNS0073W |
| C 6  | 101.8 | 14.0 | 919  | 29 | CNS048NY |
| C 7  | 100   | 14.0 | 1201 | 9  | AL581589 |
| C 8  | 99.8  | 14.0 | 1201 | 9  | AL531977 |
| C 9  | 99    | 13.8 | 1101 | 29 | CNS00EXE |
| C 10 | 96.2  | 13.5 | 753  | 13 | EX443342 |
| C 11 | 96.2  | 13.5 | 1100 | 13 | EX443342 |
| C 12 | 95.8  | 13.4 | 1097 | 29 | EX385076 |
| C 13 | 95    | 13.3 | 791  | 29 | CNS01224 |
| C 14 | 95    | 13.3 | 1101 | 29 | CNS009KS |
| C 15 | 94.6  | 13.2 | 899  | 13 | CNS00LOO |
| C 16 | 94.4  | 13.2 | 1216 | 29 | EX436853 |
| C 17 | 94.2  | 13.2 | 815  | 28 | EX436853 |
| C 18 | 94.2  | 13.2 | 1200 | 9  | EX436853 |
| C 19 | 94    | 13.1 | 1000 | 13 | AL548181 |
| C 20 | 94    | 13.1 | 1151 | 14 | EX418086 |
| C 21 | 93.8  | 13.1 | 1008 | 29 | CD501049 |
| C 22 | 93.6  | 13.1 | 884  | 29 | AG137085 |
| C 23 | 93.6  | 13.1 | 969  | 28 | CNS006UO |
| C 24 | 93.4  | 13.1 | 1172 | 29 | AG143309 |
| C 25 | 93.2  | 13.0 | 1141 | 28 | AG136828 |
| C 26 | 92.8  | 13.0 | 981  | 29 | BZ696157 |
| C 27 | 92.8  | 13.0 | 1183 | 29 | AG127518 |
| C 28 | 92.6  | 13.0 | 923  | 13 | BZ696114 |
| C 29 | 92.6  | 13.0 | 1183 | 29 | EX410248 |
| C 30 | 92.4  | 12.9 | 859  | 29 | AG136828 |
| C 31 | 92.4  | 12.9 | 880  | 29 | AG128925 |
| C 32 | 92    | 12.9 | 1137 | 28 | AG139490 |
| C 33 | 92    | 12.9 | 1225 | 29 | AG143326 |
| C 34 | 91.4  | 12.8 | 700  | 29 | CNS0166K |
| C 35 | 91.4  | 12.8 | 885  | 29 | AG127423 |
| C 36 | 91    | 12.7 | 1099 | 28 | CNS03LTM |
| C 37 | 90.8  | 12.7 | 885  | 13 | AG143360 |
| C 38 | 90.8  | 12.7 | 1081 | 29 | EX425603 |
| C 39 | 90.6  | 12.7 | 1491 | 10 | AG135328 |
| C 40 | 90.2  | 12.6 | 866  | 29 | BE882936 |
| C 41 | 90    | 12.6 | 953  | 29 | AG126308 |
| C 42 | 90    | 12.6 | 1039 | 13 | CNS008PF |
| C 43 | 90    | 12.6 | 1101 | 29 | EX378025 |
| C 44 | 89.8  | 12.6 | 896  | 29 | CNS017FC |
| C 45 | 89.8  | 12.6 | 1201 | 13 | AG141027 |
|      |       |      |      |    | BX399683 |

#### ALIGNMENTS

|            |   |        |     |        |                 |
|------------|---|--------|-----|--------|-----------------|
| RESULT 1   | CNS0071A/c  | 895 bp | DNA | linear | GSS 03-JUN-1999 |
| LOCUS      | Drosophila melanogaster genome survey sequence TET3 end of BAC #  |        |     |        |                 |
| DEFINITION | BAC14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  |        |     |        |                 |
| ACCESSION  | AL066286  |        |     |        |                 |
| VERSION    | AL066286.1  |        |     |        | GI:4945153      |
| KEYWORDS   | GSS   |        |     |        |                 |
| SOURCE     | Drosophila melanogaster (fruit fly)   |        |     |        |                 |
| ORGANISM   | Drosophila melanogaster   |        |     |        |                 |
|            | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  |        |     |        |                 |
| REFERENCE  | 1 (bases 1 to 895)  |        |     |        |                 |
| AUTHORS    | Genoscope.  |        |     |        |                 |
| TITLE      | Direct Submission   |        |     |        |                 |
| JOURNAL    | Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]  |        |     |        |                 |
| COMMENT    | Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosier in Pieter de Jong's laboratory in the Department of |        |     |        |                 |



[illegible]

```

Db      612 AAAAGHYNTYKHYTTY 595

RESULT 4
AGI27412
LOCUS   Pan troglodytes DNA, clone: PTB-138E06.R, genomic survey sequence.
DEFINITION
ACCESSION AGI27412
VERSION   AGI27412.1 GI:16656577
KEYWORDS  GSS.
SOURCE   Pan troglodytes (chimpanzee)
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 946)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS

Sequencing: M13Rev
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI.
Location/Qualifiers
1. .946
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-138E06.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
505 a 351 c 31 g 23 t 36 others

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 14.6%; Score 104.6; DB 29; Length 946;
Matches 326; Conservative 0; Mismatches 374; Indels 0; Gaps 0;

Qy 15 AACACATGCATCATCAAGATGCAACAAGCCGAGATCAAGAACACACACCCCAACATACCT 74
Db 153 AACAAACACACACCACCAANACACACCAAGGCAACAANNAAAAACCCACACCAAC 212
Qy 75 CACTCAGGATCCTCAGCTTGGATCAGTCTTCTCCAACTGTCTGAAATTTACATCAAAAC 134
Db 213 AGCAACCAAAAAACAACCAAAAAACAAAAACAACCAAGAAACAACACACACACAGAAAC 272
Qy 135 CACCACCATACTAGCTTCAACACACACGAGTCAAGTCAAACTGCAACCCACACACACT 194
Db 273 CACCCCAACAAAAAACCAAAACCAAAACCAACCAACCAACCAACCAACCAAAAAA 332
Qy 195 CAAAGCTAAACACACACCAACCAACCCAAACACAAACCCAGCAGCCCACTACAAAAACAAG 254
Db 333 AAAAACTATAACAAACACACCAACCAACCAACCCCAACCAACCAACCAACCAACCA 392
Qy 255 CCAAAACAAACCAACCAACCAACCCCAATATGATTTCTACTTCGAGTGTTTAACTTTGT 314
Db 393 AAACACACACCAACCAACCAACCCCAACCAACCCCAACCAACCAACCAACCAACCA 452

```











```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDL006YB12"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      433 a 175 c 126 g 196 t 170 others
ORIGIN
Query Match      13.5%; Score 96.2; DB 13; Length 1100;
Best Local Similarity 38.9%; Pred. No. 6.5e-06;
Matches 167; Conservative 72; Mismatches 190; Indels 0; Gaps 0;

QY 146 TAGCTTCAACACACAGGAGTCAAGTCAAACTCCACCCACACAGTCAAGCTAAA 205
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
672 TWWAWWWAAWAAAAAARAAAAAACAACAAAAAAMAAAAAAMAAAAAAMAAAAA 731
QY 206 ACACAAACACACACACACACACACACACACACACACACACACACACACAC 265
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
732 AAAAAMAAAAAAMWCAAAAAAARAAAAAACAACACACACACACACACACACAC 791
QY 266 CACCAACACACACACACACACACACACACACACACACACACACACACACAC 325
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
792 AAAMACMAAAAMCCAAAMAAAAAMWMAAMMCMCAAAAAACCCCAAAAAAACCC 851
QY 326 TATGACGACACACACACACACACACACACACACACACACACACACACACAC 385
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
852 AAAAAACACCAAMAAAAAARAAAAAACAACACACACACACACACACACACACAC 911
QY 386 GAAAGAAAAACACACACACACACACACACACACACACACACACACACACACAC 445
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
912 MAMWCAAAAAAAMWMAAAAMWMAAAAMWMAAAAMWMAAAAMWMAAAAMWMAAA 971
QY 446 TCAACCTCAACACACACACACACACACACACACACACACACACACACACACAC 505
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
972 MAAAMCCAMACCCMAAAACCMMAAAACCAAMAAACCAAAAAACCCCAAAAAAMCA 1031
QY 506 COATCAACACACACACACACACACACACACACACACACACACACACACACAC 565
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1032 AAAAAACCMCAAMCAAMWMAAAARAAAAACCAACAMWMAAMWMAAAACCCMA 1091
QY 566 ATCCAAAAAC 574
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1092 AAAMCMWCC 1100

```

```

RESULT 12
CNS01224/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN08M09 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL102202.1 GI:5613813
VERSION
GSS.
KEYWORDS
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1097)
AUTHORS
Genoscope.
DIRECT SUBMISSION
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -

```

http://www.edgp.ebi.ac.uk -- This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billand at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

```

FEATURES
source
Location/Qualifiers
1..1097
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN08M09"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/notes="end : SP6"
BASE COUNT      169 a 119 c 196 g 398 t 215 others
ORIGIN
Query Match      13.4%; Score 95.8; DB 29; Length 1097;
Best Local Similarity 32.7%; Pred. No. 7.6e-06;
Matches 199; Conservative 119; Mismatches 290; Indels 0; Gaps 0;

QY 13 CTAACAATCGAATCATACAAAGATGCAACAGCCAGATCAAGAACACAAACCCCAATAC 72
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1095 MTMCMWMTWMTCAATCTCMYCMCMCAATACMACMCMCCACMMWCAACCCCAATMCT 1036
QY 73 CTCACCTCAGGATCTCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAAATTTACATCA 132
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1035 CMCAMCMWAAACAAMACACMCMCAAMWACACACACACACACACACACACACACACACAC 976
QY 133 ACACACACACATAGCTTCAACAAACACAGGAGTCAAGTCAAAACCTGCAACCCACACAC 192
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
975 CMWCAACACAAACACACACACACACACACACACACACACACACACACACACACACAC 916
QY 193 GTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 252
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
915 AMWACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 856
QY 253 CGCCAAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 312
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
855 AHAACAHAAAAACACACACACACACACACACACACACACACACACACACACACACACAC 796
QY 313 GTACCTGTCAGCATATGAGCAACATCCACACCTGCTGGGTATCTGCAAAAGAAATACCA 372
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
795 TTTTTTTTTTTTACACACACACACACACACACACACACACACACACACACACACACAC 736
QY 373 AACAAAAACCCAGGAAAGAAACACACACACACACACACACACACACACACACACACACAC 432
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
735 WMTWACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 676
QY 433 ACCAAAAAGATCTCAAACTCAAAACCACTAAACCAAGGAGGATACCCACACACACACAC 492
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
675 AAAAMCAAMWMAACACCCCMCAACACACACACACACACACACACACACACACACACACAC 616
QY 493 ACAGAGGACCAACATCAACACACACACACACACACACACACACACACACACACACACAC 552
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
615 MAAAAAAMAAAAAMWMAACCCCMCAACACACACACACACACACACACACACACACACACAC 556
QY 553 AACACACAGGAAATCCAAACCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCA 612
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
555 TAMCMWACACACACACACACACACACACACACACACACACACACACACACACACACACAC 496
QY 613 GAAGGCAA 620
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
495 ACAACCMW 488

```

```

RESULT 13
CNS009KS
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR19F04 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
979 bp DNA linear GSS 03-JUN-1999
CNS009KS
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR19F04 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

```

```

ACCESSION      AL053801.1      GI:4935176
VERSION        GSS.
KEYWORDS
SOURCE         Drosophila melanogaster (fruit fly)
ORGANISM       Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 791)
AUTHORS
TITLE          Direct Submission
JOURNAL        Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
               BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
               - Web : www.genoscope.cns.fr)
COMMENT        Determination of this BAC-end sequence was carried out as part of a
               collaboration with the Berkeley Drosophila Genome Project (BDGP).
               The BDGP is constructing a physical map of the Drosophila
               melanogaster genome using these BACs. For further information
               please see http://www.fruitfly.org The BDGP Drosophila
               melanogaster BAC library was prepared by Kazutoyo Osoegawa and
               Aaron Mammosser in Pieter de Jong's laboratory in the Department of
               Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
               NY. The library is named RPCI-98 and was constructed by partial
               EcoRI digestion of Drosophila DNA provided by the BDGP from the
               isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
               P1 and EST libraries. A more detailed description of the library
               and how to order individual BAC clones, the entire library, or
               filters for hybridization from the BACPAC Resource Center can be
               found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES       Location/Qualifiers
               1. 791
               /organism="Drosophila melanogaster"
               /mol_type="genomic DNA"
               /db_xref="taxon:7227"
               /clone="BACR19F04"
               /clone_lib="RPCI-98"
               /note="end : T7"
BASE COUNT     464 a 105 c 67 g 78 t 77 others
ORIGIN
Query Match    13.3%; Score 95; DB 29; Length 791;
Best Local Similarity 47.1%; Pred. No. 1.e-05;
Matches 205; Conservative 43; Mismatches 178; Indels 9; Gaps 2;

QY 153 AACACACACGAGTCAAGTCAACCTGCGACCCACACAGCTCAAGCTAAGCAACAC 212
Db 153 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 92
QY 213 AACACACCAACCAACCCAGCAGCCCACTTACAAAAACAAACGCAAAACCAACAA 272
Db 93 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 152
QY 273 CAACACCA-ATAATGATTTTCACTTCGAGTGTGTTAACTTTGACCTCGCAGCATGCA 331
Db 153 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 212
QY 332 GCAACCAATCAACCTGCTGGCTATCTGCAAAAGATACCAACAAACAAACAGGAAGA 391
Db 213 ACWATAAGCAACAA-----AAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAM 264
QY 392 AAACCAACCAACGCTTACAAAAACCAACCTTCAAGACAAACCAAAAGATCTCAAC 451
Db 265 AAACCAACCAACGCTTACAAAAACCAACCTTCAAGACAAACCAAAAGATCTCAAC 324
QY 452 CTCAACCACTAAACCAAGGAGTACCCACCAAGCCCAAGCCCAAGCCCAACCATCA 511
Db 325 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 384
QY 512 ACACCAACCAACCAACATCACTACATCTGCTCACCACCAACCAACCAACCAACCA 571
Db 385 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 444
QY 572 AAUCCACCAAGTCAAA 586

```

```

Db 445 AACACACAAATATAAA 459
||| ||||| |||
CNS00LOO      1101 bp      DNA      linear      GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC;
BACR32D23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL068607
AL068607.1   GI:4958689
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES       Location/Qualifiers
               1. 1101
               /organism="Drosophila melanogaster"
               /mol_type="genomic DNA"
               /db_xref="taxon:7227"
               /clone="BACR32D23"
               /clone_lib="RPCI-98"
               /note="end : TET3"
BASE COUNT     155 a 166 c 7 g 284 t 489 others
ORIGIN
Query Match    13.3%; Score 95; DB 29; Length 1101;
Best Local Similarity 18.5%; Pred. No. 1.e-05;
Matches 110; Conservative 245; Mismatches 240; Indels 0; Gaps 0;

QY 104 TCTCCAACTGCTGGAATTAATGATCAACAAACACCAACCACTAGCTTCAACACACGAG 163
Db 455 TCTCCCTCTATTTCTBCTTTTCMCTCMCAACACACCTCTCTCTCTCTCTCTCTCTCT 514
QY 164 GAGTCAAGTCAACCTGCAACCAACCAACCAAGTCAAGCACTAAACAAACCAACCAAC 223
Db 515 CMHMMWACWMACTMMWCMWCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 574
QY 224 CACAAACCAAGCCCACTCAACAAACAAACGCAACAAACCAACCAACCAACCAACCA 283
Db 575 MAYCMYCCMYTTCMCAHHTWMAACATMMAHACHTHHTWCTWTTWTTWTTWTTWTTW 634
QY 284 ATGATTTTCACTTCGAGTGTTAATCTTGTACCTCGCAGCATATGACGAAACAATCCA 343
Db 635 MAHTMTWMAHAAHNTTWTWMMHMAHACHTHHTWCTWTTWTTWTTWTTWTTWTTW 694
QY 344 CTGCTGGGCTATCTGCAAAAGAAATACAAACCAACCAACCAACCAACCAACCAACCA 403
Db 695 MHTMTATCAAMWMAAMWMTTWTWMMWMAAMWMTTWTWMMWMAAMWMTTWTWMMWMA 754

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2003, 00:01:13 ; Search time 240.957 Seconds  
(without alignments)  
8070.528 Million cell updates/sec

Title: US-09-462-816-3

Perfect score: 715

Sequence: 1 cacaaagtcacactaacac.....gtagttattataaaaaaaa 715

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1811591 seqs, 1359896290 residues

Total number of hits satisfying chosen parameters: 3623182

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09A\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 654.8 | 91.6        | 15223  | 10 | US-09-847-173-1    |
| 2          | 445.8 | 62.3        | 696    | 10 | US-09-844-648-4    |
| 3          | 300.4 | 42.0        | 15225  | 12 | US-09-827-688-10   |
| C 4        | 85.8  | 12.0        | 367378 | 12 | US-10-312-841-2    |
| C 5        | 83.6  | 11.7        | 793    | 10 | US-09-878-574-4304 |
| C 6        | 82.6  | 11.6        | 7758   | 12 | US-10-311-455-1076 |
| C 7        | 81.4  | 11.4        | 16167  | 12 | US-10-311-455-1056 |
| C 8        | 81.4  | 11.4        | 16167  | 12 | US-10-240-488-82   |
| C 9        | 80    | 11.2        | 6668   | 12 | US-10-311-455-1670 |
| C 10       | 78.8  | 11.0        | 1972   | 12 | US-10-017-161-1981 |
| C 11       | 76.6  | 10.7        | 576    | 10 | US-09-878-574-4296 |
| C 12       | 76.4  | 10.7        | 529    | 10 | US-09-983-965-2109 |
| C 13       | 75.6  | 10.6        | 1030   | 10 | US-09-878-574-4306 |
| C 14       | 75    | 10.5        | 577    | 10 | US-09-878-574-4312 |
| C 15       | 74.8  | 10.5        | 6089   | 12 | US-10-311-455-675  |
| C 16       | 74.8  | 10.5        | 11622  | 12 | US-10-311-455-648  |

|      |      |      |       |    |                     |                    |
|------|------|------|-------|----|---------------------|--------------------|
| C 17 | 74   | 10.3 | 639   | 10 | US-09-878-574-4316  | Sequence 4316, Ap  |
| C 18 | 73.4 | 10.3 | 6668  | 12 | US-10-311-455-1669  | Sequence 1669, Ap  |
| C 19 | 72.8 | 10.2 | 1083  | 12 | US-10-029-386-20723 | Sequence 20723, A  |
| C 20 | 72.8 | 10.2 | 6048  | 12 | US-10-311-455-2004  | Sequence 2004, Ap  |
| C 21 | 72.2 | 10.1 | 13606 | 12 | US-10-311-455-1793  | Sequence 1793, Ap  |
| C 22 | 72.2 | 10.1 | 13606 | 12 | US-10-240-453-187   | Sequence 187, App  |
| C 23 | 72.2 | 10.1 | 13606 | 14 | US-10-239-676-165   | Sequence 165, App  |
| C 24 | 71.8 | 10.0 | 2543  | 12 | US-10-029-386-20536 | Sequence 20536, A  |
| C 25 | 71.8 | 10.0 | 3163  | 12 | US-10-017-161-1857  | Sequence 1857, Ap  |
| C 26 | 71.8 | 10.0 | 14006 | 12 | US-10-311-455-1931  | Sequence 1931, Ap  |
| C 27 | 71.4 | 10.0 | 1075  | 9  | US-09-864-761-19241 | Sequence 19241, A  |
| C 28 | 71.4 | 10.0 | 1403  | 9  | US-09-864-761-19241 | Sequence 2513, Ap  |
| C 29 | 71.2 | 10.0 | 1635  | 9  | US-09-864-761-20241 | Sequence 20241, A  |
| C 30 | 71.2 | 10.0 | 1973  | 9  | US-09-864-761-3471  | Sequence 3471, Ap  |
| C 31 | 71   | 9.9  | 13732 | 12 | US-10-311-455-1793  | Sequence 1793, Ap  |
| C 32 | 70.8 | 9.9  | 11996 | 12 | US-10-240-485-45    | Sequence 45, Appl  |
| C 33 | 70.6 | 9.9  | 8147  | 12 | US-10-311-455-401   | Sequence 401, Appl |
| C 34 | 70.6 | 9.9  | 9415  | 12 | US-10-311-455-268   | Sequence 268, App  |
| C 35 | 70.4 | 9.8  | 6270  | 12 | US-10-311-455-378   | Sequence 378, App  |
| C 36 | 70.4 | 9.8  | 6270  | 12 | US-10-240-453-20    | Sequence 20, Appl  |
| C 37 | 70.4 | 9.8  | 6270  | 14 | US-10-239-676-12    | Sequence 12, Appl  |
| C 38 | 70.4 | 9.8  | 9539  | 12 | US-10-240-453-54    | Sequence 54, Appl  |
| C 39 | 70.4 | 9.8  | 9539  | 14 | US-10-239-676-52    | Sequence 52, Appl  |
| C 40 | 70.2 | 9.8  | 10467 | 12 | US-10-240-453-327   | Sequence 327, App  |
| C 41 | 70.2 | 9.8  | 17869 | 12 | US-10-311-455-78    | Sequence 78, Appl  |
| C 42 | 70   | 9.8  | 576   | 9  | US-09-864-761-26582 | Sequence 26582, A  |
| C 43 | 70   | 9.8  | 6065  | 12 | US-10-311-455-477   | Sequence 477, App  |
| C 44 | 69.6 | 9.7  | 5774  | 12 | US-10-311-455-1136  | Sequence 1136, Ap  |
| C 45 | 69.6 | 9.7  | 9219  | 12 | US-10-311-455-2400  | Sequence 2400, Ap  |

#### ALIGNMENTS

#### RESULT 1

US-09-847-173-1  
; Sequence 1, Application US/09847173  
; Publication No. US2002018228A1  
; GENERAL INFORMATION:  
; APPLICANT: Collins, Peter L.  
; TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY  
; SYNCTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/847,173  
; FILING DATE: 03-May-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/720,132  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; TELEPHONE: 206-467-9600  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15223 base pairs  
; TYPE: nucleic acid



```
;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-847-173-1

Query Match      91.6%; Score 654.8; DB 10; Length 15223;
Best Local Similarity 94.8%; Pred. No. 4e-164;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACCACTGCAATCATACAAAGATGCAACAAAGCCAGATCAAGAACACA 60
DB
4887 CACAAAGTCACACCAACCACTGCAATCATACAAAGATGCAACAAAGCCAGATCAAGAACACA 4946

QY 61 ACCCCAAACATCCTCACTCAGATCCTCAGCTTGGAAATCAGTCTTCCAACTCTGCTGAA 120
DB
4947 ACCCCAAACATCCTCACTCAGATCCTCAGCTTGGAAATCAGTCTTCCAACTCTGCTGAA 5006

QY 121 ATTACATCAAAACACCAACCATCTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 180
DB
5007 ATTACATCAAAACACCAACCATCTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 5066

QY 181 CAACCCAAACAGTCAGATCAAGATCAAAACCAACAAACCCCAACCAACCAACCAAGGCC 240
DB
5067 CAATCCAAACAGTCAGATCAAGATCAAAACCAACCAACCAACCAACCAAGGCC 5126

QY 241 ACTCAAAACAAACGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB
5127 ACCCAAAACAAACGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186

QY 301 GTGTTTAACTTTGTACCTGAGCATATGTCAGCAACATCAACCTGCTGGGCTATCTGC 360
DB
5187 GTGTTTAACTTTGTACCTGAGCATATGTCAGCAACATCAACCTGCTGGGCTATCTGC 5246

QY 361 AAAAGAATACCAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB
5247 AAAAGAATACCAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5306

QY 421 ACCTTCAGACACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB
5307 ACCCTTCAGACACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5366

QY 481 ACCCAAGGCCACAGAGAGCAACCATCAACACCAACCAACCAACCAACCAACCAACCAACCA 540
DB
5367 ACCCAAGGCCACAGAGAGCAACCATCAACACCAACCAACCAACCAACCAACCAACCAACCA 5426

QY 541 CTGCTTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB
5427 CTGCTTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5486

QY 601 TCAACCTCTCGAAGGCAATTAAGCCCTTCTCAAGTCTTCAACCAACCAACCAACCAACCAACCA 660
DB
5487 TCAACCTCTCGAAGGCAATTAAGCCCTTCTCAAGTCTTCAACCAACCAACCAACCAACCAACCA 5546

QY 661 TCAACACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 714
DB
5547 TCAACACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5600
```

## RESULT 2

```
US-09-844-645-4
; Sequence 4, Application US/09844645
; Patent No. US20020102242A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
;              McDaniel, Larry S.
;              Curiel, David T.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
; ADMINISTERING PNEUMOCOCCAL DNA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
```

```
;
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/844,645
; FILING DATE: 27-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,505
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-844-645-4
```

```
Query Match      62.3%; Score 445.8; DB 10; Length 696;
Best Local Similarity 95.3%; Pred. No. 5.8e-109;
Matches 470; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 1 CACAAAGTCACACTAACCACTGCAATCATACAAAGATGCAACAAAGCCAGATCAAGAACACA 60
DB
199 CACAAAGTCACACCAACCAACCACTGCAATCATACAAAGATGCAACAAAGCCAGATCAAGAACACA 258

QY 61 ACCCCAAACATCCTCACTCAGATCCTCAGCTTGGAAATCAGTCTTCCAACTCTGCTGAA 120
DB
259 ACCCCAAACATCCTCACTCAGATCCTCAGCTTGGAAATCAGTCTTCCAACTCTGCTGAA 318

QY 121 ATTACATCAAAACCAACCAACCACTAGCTTCAACCAACCAAGGAGTCAAGTCAAACTG 180
DB
319 ATTACATCAAAACCAACCAACCACTAGCTTCAACCAACCAAGGAGTCAAGTCAAACTG 378

QY 181 CAACCCCAACAGTCAGATCAAGATCAAAACCAACCAACCAACCAACCAACCAAGGCC 239
DB
379 CAATCCCAACAGTCAGATCAAGATCAAAACCAACCAACCAACCAACCAAGGCC 438

QY 240 CACTTCAAAACCAACGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 299
DB
439 CACCAAAACCAACGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498

QY 300 AGTGTTTAACTTTGTACCTGCGAGCATATGCGACCAACCAATCCAACTCTGCTGGGCTATCTG 359
DB
499 AGTGTTCACCTTTGTACCTGCGAGCATATGCGACCAACCAATCCAACTCTGCTGGGCTATCTG 558

QY 360 CAAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 419
DB
559 CAAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618

QY 420 AACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 479
DB
619 AACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678

QY 480 CACCACCAAGGCC 492
DB
679 CACCACCAAGGCC 691
```

## RESULT 3

US-09-827-688-10  
; Sequence 10, Application US/09827688  
; Publication No. US20030165476A1  
; GENERAL INFORMATION:  
; APPLICANT: ORSON, FRANK  
; APPLICANT: KINSEY, BERNIA  
; APPLICANT: BHOAGAL, BALBIR  
; TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION  
; FILE REFERENCE: P01949US1/10004014  
; CURRENT APPLICATION NUMBER: US/09/827,688  
; CURRENT FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: 2001-04-06  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; TYPE: DNA  
; ORGANISM: RSV  
US-09-827-688-10

Query Match 42.0%; Score 300.4; DB 12; Length 15225;  
Best Local Similarity 64.6%; Pred. No. 1.5e-69;  
Matches 464; Conservative 0; Mismatches 251; Indels 3; Gaps 1;

|    |      |  |      |
|----|------|--|------|
| QY | 1    | CACAAAGTTCACACTAACACTGCAATCATACAGATGCAACAGCCAGCCAGATCAAGAACACA   | 60   |
| DB | 4888 | CACAAAGTTACACTAACAGGTTCACAGTTCACACATATAAAACCCACTGAAAAAAC         | 4947 |
| QY | 61   | ACCCCAACATCTCTACTCAGGATCTCTAGCTTTGGAATCAGCTTTCTCCATCTGTCTGAA     | 120  |
| DB | 4948 | ATCACCACCTTACTTACTTACTCAAGTCCACACAGAAAGGTTAGTCATCTCCAAACACCTACA  | 5007 |
| QY | 121  | ATTACATCAACACCACTACCTACTAGCTTTCAACACACGAGGAGTCAAGTCAAACTTG       | 180  |
| DB | 5008 | ACCACATTCACCAATCCACACAAATTCAGCCACACATCTACCCCAACAAAGTCAGAAACA     | 5067 |
| QY | 181  | CAACCCACACAGCTCAAGACTTAAACACACAAACCAACCAACCAACCAACCAACCAACCA     | 240  |
| DB | 5068 | CACCAACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA    | 5127 |
| QY | 241  | ACTACAAACACCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA    | 300  |
| DB | 5128 | AGCACAAACACCGCTTAAACCAATCCACCAACCAACCAACCAACCAACCAACCAACCAACCA   | 5187 |
| QY | 301  | GTGTTTAACTTTGTACCTTGCAGCATATGACGACCAACATCCACCTGCTGGGCTATCTGC     | 360  |
| DB | 5188 | GTGTTTCACTTGTCTTCTGTATATATGTGGCAACCAATCAACTTTGCAATCTCATCTGT      | 5247 |
| QY | 361  | AAAAGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA     | 420  |
| DB | 5248 | AAAACATACCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  | 5307 |
| QY | 421  | ACCTTCAAG---ACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA   | 477  |
| DB | 5308 | ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  | 5367 |
| QY | 478  | CCACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA   | 537  |
| DB | 5368 | ACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  | 5427 |
| QY | 538  | ACACTGTCTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA   | 597  |
| DB | 5428 | TCAATTCATCTGTCTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA | 5487 |
| QY | 598  | CACTCAACTCTCTCGAAGGCAATCTAAGCCCTTTCTCAAGTCTCCACACATCTCCGAGCAC    | 657  |
| DB | 5488 | CACTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA   | 5547 |
| QY | 658  | CAATCAACCAACCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA | 715  |

Db 5548 TCTACATCAAAATTCACCCAAATACCAATACATGCTTAGTTATTTCAAAAACACTACA 5605

## RESULT 4

US-10-312-841-2/c  
; Sequence 2, Application US/10312841  
; Publication No. US20030186277A1  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des  
; FILE REFERENCE: E01/1208/WO  
; CURRENT APPLICATION NUMBER: US/10/312,841  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 2  
; SEQ ID NO 2  
; LENGTH: 3673778  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (379615)  
US-10-312-841-2

Query Match 12.0%; Score 85.8; DB 12; Length 3673778;  
Best Local Similarity 47.0%; Pred. No. 3.1e-11;  
Matches 334; Conservative 0; Mismatches 372; Indels 5; Gaps 2;

|    |         |   |         |
|----|---------|---|---------|
| QY | 9       | CACACTAACCAACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACACACACCCCAAC  | 68      |
| DB | 2435804 | CTCAACAAACCTTACAAAACACGAAAAAACTCCACTAACATATAAAACACACACACC       | 2435745 |
| QY | 69      | ATACCTCACTCAGGATCTCTCAGCTTGAATCAGCTTCTCCATCTGTCTGAAATTCATC      | 128     |
| DB | 2435744 | ATCCCGCAAAACCTTACAAAAATAAAAAAACAACCAATATAAAACACACACACATC        | 2435685 |
| QY | 129     | ACAAACCCACCCATCTAGCTTCAACACACAGGAGTCAAGTCAAACTGCAACCCAC         | 188     |
| DB | 2435684 | CCACACAAACCTTACAAAAATAAAAAAATAACCAACGAAAAAACCACTATTCC           | 2435625 |
| QY | 189     | AACAGTCAAGACTAAAAACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA   | 248     |
| DB | 2435624 | AACAAACCTTACAAAAATAAAAAAATAACCAACCAATATAAAACCACTATCCCAAC        | 2435565 |
| QY | 249     | ACAAGCCAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  | 308     |
| DB | 2435564 | AAAC-CTACAAAAATAAACAAAAAACCCCAATTTACCAATAAAAAAACCACTATCCC       | 2435506 |
| QY | 309     | CTTTGTACCTGCAGCATATGACGACCAACATCCAACTGTCTGGCTATCTGCAAAAGAT      | 368     |
| DB | 2435505 | CAACAAACCTTACAAAAACACGAAAAAAACCCCACTAACCAATAAAAAACACCACTAT      | 2435446 |
| QY | 369     | ACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA    | 424     |
| DB | 2435445 | CCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA    | 2435386 |
| QY | 425     | TCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA | 484     |
| DB | 2435385 | CAACAAACCTTACAAAAATAAAAAAATAACCAACCAATAAAAAACCACTATCCCAAC       | 2435326 |
| QY | 485     | CCAGCCCAACGAGGAGCCACCATCAACACCAACCAACCAACCAACCAACCAACCAACCAACCA | 544     |
| DB | 2435325 | CACACCTTACAAAAATAAAAAAATAACCAACCAATAAAAAAACCACTATCCCAAC         | 2435266 |
| QY | 545     | TCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  | 604     |
| DB | 2435265 | TACAAACCTTACAAAAATAAAAAAATAACCAACCAATAAAAAAACCACTATCCCAAC       | 2435206 |
| QY | 605     | CCTCTCCGAGGCAATCTAAGCCCTTTCTCAAGTCTCCACACATCTCCGAGCAACCCATCAC   | 664     |
| DB | 2435205 | CCTCAACAAACCTTACAAAAACACGAAAAAACTCCCACTAACCAATAAAAAACCACTAT     | 2435146 |



Db 2057 ACTCCCAACACCAAAAAAACAACCTTCTTAAACACCAAAAAAAC 2005

RESULT 7

US-10-311-455-1056/c

; Sequence 1056, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311.455

; PRIOR FILING DATE: 2002-12-16

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 1056

; LENGTH: 16167

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; NAME/KEY: unsure

; LOCATION: 4003, 5529, 5534, 5553, 5570, 5573, 5577..5578, 5592, 5732

; OTHER INFORMATION: n is a or g or c or t

; NAME/KEY: unsure

; LOCATION: 5837..5838

; OTHER INFORMATION: n is a or g or c or t

US-10-311-455-1056

Query Match 11.4%; Score 81.4; DB 12; Length 16167;

Best Local Similarity 46.6%; Pred. No. 4e-11;

Matches 235; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

QY 109 AATCTGTCTGAAATTACATCACAAACCCACCACCATAGCTTCAACAAACACCCAGGATC 168

Db 5866 ATTCTTACTTTTCTCACCCAAACGNNNAACAAACGAAAAAATAAAAAAATAAAAAA 5807

QY 169 AGTCAAACTGCAACCCACACAGTCAAGCTAATAAACAACAACCAACCCCAACACAA 228

Db 5806 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5747

QY 229 CCAGCAAGCCCTACAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCA 288

Db 5746 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5687

QY 289 TTTCACTTGAAGTGTGTAATCTTGTACCTGTGAGCATATGAGCAATATGAT 348

Db 5686 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5627

QY 349 TGGGCTATCTGCAAAAGATATACCAACCAACCAACCAACCAACCAACCAACCAAC 408

Db 5626 CGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5567

QY 409 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 468

Db 5566 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5507

QY 469 AAGGAGTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 528

Db 5506 AACCTTTTACCAAAACCAATTTACATCATCCCTTAACCTTAACCTTAACCTTAAC 5447

QY 529 ATCAACATACATGCTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 588

Db 5446 TCTCCCAATTCCTTCAAACTTCTTATCACACCTTAATAAAAAAATAAAAAA 5387

QY 589 GAAACCTTCCACTCAACCTCTCTCC 612

Db 5386 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5363

RESULT 8

US-10-240-485-82/c

; Sequence 82, Application US/10240485

; Publication No. US20030148327A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with

; TITLE OF INVENTION: Metastasis

; FILE REFERENCE: 5013.1007

; CURRENT APPLICATION NUMBER: US/10/240.485

; PRIOR FILING DATE: 2002-10-02

; PRIOR APPLICATION NUMBER: PCT/EP01/03970

; PRIOR FILING DATE: 2001-04-06

; PRIOR APPLICATION NUMBER: DE 10019058.8

; PRIOR FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: DE 10019173.8

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 202

; SEQ ID NO 82

; LENGTH: 16167

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; NAME/KEY: unsure

; LOCATION: 4003, 5529, 5534, 5553, 5570, 5573, 5577..5578, 5592, 5732)

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (5837..)

US-10-240-485-82

Query Match 11.4%; Score 81.4; DB 12; Length 16167;

Best Local Similarity 46.6%; Pred. No. 4e-11;

Matches 235; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

QY 109 AATCTGTGAAATTACATCACAAACCCACCACCATAGCTTCAACAAACACCCAGGATC 168

Db 5866 ATTCTTACTTTTCTCACCCAAACGNNNAACAAACGAAAAAATAAAAAAATAAAAAA 5807

QY 169 AAGTCAAACTGCAACCCACACAGTCAAGCTAATAAACAACAACCAACCAACCAACCA 228

Db 5806 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5747

QY 229 CCAGCAAGCCCTACAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCA 288

Db 5746 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5687

QY 289 TTTCACTTGAAGTGTGTAATCTTGTACCTGTGAGCATATGAGCAATATGAT 348

Db 5686 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5627

QY 349 TGGGCTATCTGCAAAAGATATACCAACCAACCAACCAACCAACCAACCAACCAAC 408

Db 5626 CGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5567

QY 409 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 468

Db 5566 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5507







## RESULT 13

US-09-878-574-4306/c  
; Sequence 4306, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Thompson, Michael D.  
; APPLICANT: La Rosa, Thomas J.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1998-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 4306  
; LENGTH: 1030  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1030)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-D11  
US-09-878-574-4306

Query Match 10.6%; Score 75.6; DB 10; Length 1030;  
Best Local Similarity 41.0%; Pred. No. 4.1e-10;  
Matches 283; Conservative 0; Mismatches 397; Indels 11; Gaps 2;  
  
QY 1 CACAAAGTCACACTAACCTCACTCAGGATCCTCAGTTGGAATCAGCTTCTCCAAATCTGCTGAA 60  
DB 789 GACAAAC 730  
  
QY 61 ACCCAACACATCCTCCTCAGGATCCTCAGTTGGAATCAGCTTCTCCAAATCTGCTGAA 120  
DB 729 AACAAAC 671  
  
QY 121 ATTACATCAAC 180  
DB 670 -----NACAAAC 618  
  
QY 181 CAACCCAC 240  
DB 617 CANANAC 558  
  
QY 241 ACTACAAAC 300  
DB 557 ANCNANANAC 498  
  
QY 301 GTGTTTAACTTTGACCTTGAGCATATGACGACACATCTCAACCTGCTGGGCTATCTGC 360  
DB 497 CCCCCAC 438  
  
QY 361 AAAAGAAATACCAAC 420  
DB 437 ACNAC 378  
  
QY 421 ACCTTCAAGAC 480  
DB 377 ACAAAC 321  
  
QY 481 ACCACAGCCAC 540  
DB 320 ANAAAC 261  
  
QY 541 CTGCTTCAAC 600  
DB 260 CNACNAAAGCNAC 201  
  
QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGCTCTCCACACACATCCGAGCACCCA 660

## RESULT 14

US-09-878-574-4312/c  
; Sequence 4312, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Thompson, Michael D.  
; APPLICANT: La Rosa, Thomas J.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 4312  
; LENGTH: 577  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(577)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-G11  
US-09-878-574-4312

Query Match 10.5%; Score 75; DB 10; Length 577;  
Best Local Similarity 50.0%; Pred. No. 4.6e-10;  
Matches 232; Conservative 0; Mismatches 230; Indels 2; Gaps 2;  
  
QY 131 AAACCCACACCATCTAGCTTCAACCAACACACACACACACACACACACACACACACAC 190  
DB 533 AAAC 474  
  
QY 191 CAGTCAAGACTTAAAC 250  
DB 473 AACCAAC 414  
  
QY 251 AACCCCAAC 309  
DB 413 CACAAC 354  
  
QY 310 TTTGTACCTGCGAGCATATGAGCAACATCTCAACCTGCTGGGCTATCTGCAAAAGATA 369  
DB 353 CAAAAACCCCAACAA 294  
  
QY 370 CCAAC 428  
DB 293 CCAAC 234  
  
QY 429 GAC 488  
DB 233 AANAAACCCCAAC 174  
  
QY 489 GCCCACAAGAGAGCAACCATCAACACACACACACACACACACACACACACACACACAC 548  
DB 173 CACCCCAAC 114  
  
QY 549 CAAC 592  
DB 113 CAAC 70

## RESULT 15

US-10-311-455-675/c



Search completed: October 30, 2003, 03:23:54  
Job time : 246.957 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 21:37:29 ; Search time 55.5382 Seconds  
(without alignments)  
5682.373 Million cell updates/sec

Title: US-09-462-816-3  
Perfect score: 715  
Sequence: 1 cacaaagtcacactaacac.....gtagtattataaaaaaaaaa 715

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 715   | 100.0       | 920    | 2     | US-08-467-963C-7  |
| 2          | 715   | 100.0       | 920    | 2     | US-08-838-189D-7  |
| 3          | 715   | 100.0       | 920    | 3     | US-08-852-344D-7  |
| 4          | 715   | 100.0       | 920    | 3     | US-08-344-639E-7  |
| 5          | 715   | 100.0       | 920    | 3     | US-08-467-969A-7  |
| 6          | 715   | 100.0       | 920    | 3     | US-08-467-961A-7  |
| 7          | 715   | 100.0       | 920    | 3     | US-08-001-554A-7  |
| 8          | 696   | 97.3        | 894    | 2     | US-08-467-963C-28 |
| 9          | 696   | 97.3        | 894    | 2     | US-08-838-189D-28 |
| 10         | 696   | 97.3        | 894    | 3     | US-08-852-344D-28 |
| 11         | 696   | 97.3        | 894    | 3     | US-08-344-639E-28 |
| 12         | 654.8 | 91.6        | 15222  | 2     | US-08-801-898A-23 |
| 13         | 654.8 | 91.6        | 15222  | 2     | US-08-962-690-12  |
| 14         | 654.8 | 91.6        | 15223  | 2     | US-08-862-403A-1  |
| 15         | 654.8 | 91.6        | 15223  | 3     | US-08-720-132-1   |
| 16         | 300.4 | 42.0        | 15225  | 2     | US-08-892-403A-2  |
| 17         | 212.8 | 29.8        | 1050   | 4     | US-08-836-501-76  |
| 18         | 212.8 | 29.8        | 1050   | 4     | US-09-626-830-76  |
| 19         | 210.8 | 29.5        | 303    | 3     | US-08-721-979A-1  |
| 20         | 210.8 | 29.5        | 303    | 3     | US-08-836-504A-1  |
| 21         | 210.8 | 29.5        | 303    | 3     | US-08-836-501-1   |
| 22         | 210.8 | 29.5        | 303    | 4     | US-09-654-289-1   |
| 23         | 210.8 | 29.5        | 303    | 4     | US-09-582-876-1   |
| 24         | 210.8 | 29.5        | 303    | 4     | US-09-626-830-1   |
| 25         | 209.6 | 29.3        | 1071   | 3     | US-08-836-501-77  |
| 26         | 209.6 | 29.3        | 1071   | 4     | US-08-626-830-77  |
| 27         | 207.6 | 29.0        | 303    | 3     | US-08-721-979A-3  |

28 207.6 29.0 303 3 US-08-836-504A-3 Sequence 3, Appli  
29 207.6 29.0 303 3 US-08-836-501-3 Sequence 3, Appli  
30 207.6 29.0 303 4 US-09-654-289-3 Sequence 3, Appli  
31 207.6 29.0 303 4 US-09-582-876-3 Sequence 3, Appli  
32 207.6 29.0 303 4 US-09-626-830-3 Sequence 3, Appli  
33 201.2 28.1 303 3 US-08-721-979A-14 Sequence 14, Appli  
34 201.2 28.1 303 3 US-08-836-501-14 Sequence 14, Appli  
35 201.2 28.1 303 4 US-09-654-289-14 Sequence 14, Appli  
36 201.2 28.1 303 4 US-09-582-876-14 Sequence 14, Appli  
37 201.2 28.1 303 4 US-09-626-830-14 Sequence 14, Appli  
38 128.6 18.0 183 3 US-08-721-979A-23 Sequence 23, Appli  
39 128.6 18.0 183 3 US-08-836-501-23 Sequence 23, Appli  
40 128.6 18.0 183 4 US-09-654-289-23 Sequence 23, Appli  
41 128.6 18.0 183 4 US-09-582-876-23 Sequence 23, Appli  
42 128.6 18.0 183 4 US-09-626-830-23 Sequence 23, Appli  
43 126 17.6 303 3 US-08-721-979A-2 Sequence 2, Appli  
44 126 17.6 303 3 US-08-836-504A-2 Sequence 2, Appli  
45 126 17.6 303 3 US-08-836-501-2 Sequence 2, Appli

#### ALIGNMENTS

##### RESULT 1

US-08-467-963C-7

; Sequence 7, Application US/08467963C

; Patent No. 5968776

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michel H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHYN, Mary E

; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST

; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6th Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1K7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER: US/08/467,963C

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/838,189

; FILING DATE: 16-APR-1997

; APPLICATION NUMBER: US 08/001,554

; FILING DATE: 06-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9200117.1

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 920 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-467-963C-7

```
Query Match      100.0%; Score 715; DB 2; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTCACTCAATCATCAAGATGCAACAAAGCCAGATCAAGAACACA 60
DB 206 CACAAAGTCACACTAACTCAATCATCAAGATGCAACAAAGCCAGATCAAGAACACA 265

QY 61 ACCCAAACATACCTCACTCAGGATCCTCAGTTGGAATCAGCTTCTCCAACTCTGTGAA 120
DB 266 ACCCAAACATACCTCACTCAGGATCCTCAGTTGGAATCAGCTTCTCCAACTCTGTGAA 325

QY 121 ATTACATCAAAACCAACCACTACTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 180
DB 326 ATTACATCAAAACCAACCACTACTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 385

QY 181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAAC 240
DB 386 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAAC 445

QY 241 ACTTCAAAACAAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
DB 446 ACTTCAAAACAAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 505

QY 301 GTGTTTAACTTTGTACCTTGCAGCATATGAGCAACCAATCCAACTCTGGGCTATCTGC 360
DB 506 GTGTTTAACTTTGTACCTTGCAGCATATGAGCAACCAATCCAACTCTGGGCTATCTGC 565

QY 361 AAAAGATAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAGATAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625

QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685

QY 481 ACACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
DB 686 ACACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 745

QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
DB 746 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 805

QY 601 TCAACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCCGAGCACCA 660
DB 806 TCAACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCCGAGCACCA 865

QY 661 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB 866 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920
```

## RESULT 2

```
US-08-838-189D-7
; Sequence 7, Application US/08038189D
; Patent No. 598169
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: ENASHVYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,189D
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-838-189D-7

Query Match      100.0%; Score 715; DB 2; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTCAATCATCAAGATGCAACAAAGCCAGATCAAGAACACA 60
DB 206 CACAAAGTCACACTAACTCAATCATCAAGATGCAACAAAGCCAGATCAAGAACACA 265

QY 61 ACCCAAACATACCTCACTCAGGATCCTCAGTTGGAATCAGCTTCTCCAACTCTGTGAA 120
DB 266 ACCCAAACATACCTCACTCAGGATCCTCAGTTGGAATCAGCTTCTCCAACTCTGTGAA 325

QY 121 ATTACATCAAAACCAACCACTACTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 180
DB 326 ATTACATCAAAACCAACCACTACTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 385

QY 181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAAC 240
DB 386 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAAC 445

QY 241 ACTTCAAAACAAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
DB 446 ACTTCAAAACAAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 505

QY 301 GTGTTTAACTTTGTACCTTGCAGCATATGAGCAACCAATCCAACTCTGGGCTATCTGC 360
DB 506 GTGTTTAACTTTGTACCTTGCAGCATATGAGCAACCAATCCAACTCTGGGCTATCTGC 565

QY 361 AAAAGATAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAGATAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625

QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685

QY 481 ACACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
DB 686 ACACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 745

QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
DB 746 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 805

QY 601 TCAACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCCGAGCACCA 660
DB 806 TCAACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCCGAGCACCA 865

QY 661 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB 866 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920
```

QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGCACCA 660  
DB 806 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGCACCA 865  
QY 661 TCAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGCACCA 715  
DB 866 TCAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGCACCA 920

RESULT 3  
US-08-852-344D-7  
; Sequence 7, Application US/08852344D  
; Patent No. 6017539  
; GENERAL INFORMATION:  
; APPLICANT: KLEIN, Michel H  
; APPLICANT: DU, Run-Pan  
; APPLICANT: EWASYSHYN, Mary E  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION  
; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY  
; TITLE OF INVENTION: SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/852,344D  
; FILING DATE: 07-MAY-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/344,639  
; FILING DATE: 14-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-688 MTS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 920 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-852-344D-7

Query Match 100.0%; Score 715; DB 3; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.9e-185;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTCACACTAACCTGCAATCATACAGATCAACAGCCAGATCAAGACACA 60  
DB 206 CACAAGTCACACTAACCTGCAATCATACAGATCAACAGCCAGATCAAGACACA 265  
QY 61 ACCCAACATACCTCACTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAATCTCTGAA 120  
DB 266 ACCCAACATACCTCACTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAATCTCTGAA 325  
QY 121 ATTATATCAACAACCAACCATAGCTTCAACACACAGGAGTCAGATCAAACTG 180

DB 326 ATTATATCAACAACCAACCATAGCTTCAACACACAGGAGTCAGATCAAACTG 385  
QY 181 CAAACCAACAGTCAGAGCTTAAACCAACCAACCAACCAACCAACCAACCAACCAAC 240  
DB 386 CAAACCAACAGTCAGAGCTTAAACCAACCAACCAACCAACCAACCAACCAACCAAC 445  
QY 241 ACTACAAACCAACCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300  
DB 446 ACTACAAACCAACCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 505  
QY 301 GTGTTTAACTTTGTACCTGTCAGCATATGAGCAACCAACCAACCAACCAACCAACCAAC 360  
DB 506 GTGTTTAACTTTGTACCTGTCAGCATATGAGCAACCAACCAACCAACCAACCAACCAAC 565  
QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
DB 566 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
QY 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480  
DB 626 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 685  
QY 481 ACCCAACGCCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCAAC 540  
DB 686 ACCCAACGCCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCAAC 745  
QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600  
DB 746 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 805  
QY 601 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGCACCA 660  
DB 806 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGCACCA 865  
QY 661 TCACAACCTCATCTTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 715  
DB 866 TCACAACCTCATCTTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 920

RESULT 4  
US-08-344-639E-7  
; Sequence 7, Application US/08344639E  
; Patent No. 6033668  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasysyn, Mary E  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS  
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
; TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/344,639E  
; FILING DATE: 14-NOV-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1

```
/ FILING DATE: 06-JAN-1992
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stewart, Michael I
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-391 MIS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/ TELEFAX: (416) 595-1163
/ TELEX: 065-24567 SIMBAS
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 920 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-344-639E-7

Query Match
Best Local Similarity 100.0%; Score 715; DB 3; Length 920;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACCTGCAATCATACAAAGTCAAGCAAGCCAGATCAAGAACACA 60
DB |
DB 206 CACAAAGTCACACTAACCTGCAATCATACAAAGTCAAGCAAGCCAGATCAAGAACACA 265
QY 61 ACCCCAAATACCTCACTCAGGATCCTCAGCTTGAATCAGCTTCCCAATCTGCTGAA 120
DB |
DB 266 ACCCCAAATACCTCACTCAGGATCCTCAGCTTGAATCAGCTTCCCAATCTGCTGAA 325
QY 121 ATTACATCAAAACCCACCACCATCTAGCTTCAACCAACACAGGAGTCAAGTCAAACTG 180
DB |
DB 326 ATTACATCAAAACCCACCACCATCTAGCTTCAACCAACACAGGAGTCAAGTCAAACTG 385
QY 181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 240
DB |
DB 386 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCC 445
QY 241 ACTACAAACAGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB |
DB 446 ACTACAAACAGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTTAACTTGTACCTGAGCATATGTCAGCAACCAATCCAACTGCTGGCTATCTGC 360
DB |
DB 506 GTGTTTAACTTGTACCTGAGCATATGTCAGCAACCAATCCAACTGCTGGCTATCTGC 565
QY 361 AAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB |
DB 566 AAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCTTCAGAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB |
DB 626 ACCTTCAGAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685
QY 481 ACCCAAGGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
DB |
DB 686 ACCCAAGGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB |
DB 746 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805
QY 601 TCAACCTCTCCGAGGCAATCTAGCCCTTCTCAGCTTCCAGCAACCAATCCGAGCACCA 660
DB |
DB 806 TCAACCTCTCCGAGGCAATCTAGCCCTTCTCAGCTTCCAGCAACCAATCCGAGCACCA 865
QY 661 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB |
DB 866 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920
```

RESULT 5

```
US-08-467-969A-7
; Sequence 7, Application US/08467969A
; Patent No. 6168786
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysbyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,969A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-475 MIS:bh
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-467-969A-7
```

```
Query Match 100.0%; Score 715; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACCTGCAATCATACAAAGTCAAGCAAGCCAGATCAAGAACACA 60
DB |
DB 206 CACAAAGTCACACTAACCTGCAATCATACAAAGTCAAGCAAGCCAGATCAAGAACACA 265
QY 61 ACCCCAAATACCTCACTCAGGATCCTCAGCTTGAATCAGCTTCCCAATCTGCTGAA 120
DB |
DB 266 ACCCCAAATACCTCACTCAGGATCCTCAGCTTGAATCAGCTTCCCAATCTGCTGAA 325
QY 121 ATTACATCAAAACCCACCACCATCTAGCTTCAACCAACACAGGAGTCAAGTCAAACTG 180
DB |
DB 326 ATTACATCAAAACCCACCACCATCTAGCTTCAACCAACACAGGAGTCAAGTCAAACTG 385
QY 181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCC 240
DB |
DB 386 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCC 445
QY 241 ACTACAAACAGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 300
DB |
DB 446 ACTACAAACAGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 505
```



```
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/001.554A
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-286
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-001-554A-7

Query Match 100.0%; Score 715; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CACAAAGTCACACTAACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60
Db |
Qy 206 CACAAAGTCACACTAACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 265
Db |
Qy 61 ACCCAACATACCTCCTCAGATCCTCAGTTGGAATCAGTTCTTCAATCTGTGAA 120
Db |
Qy 266 ACCCAACATACCTCCTCAGATCCTCAGTTGGAATCAGTTCTTCAATCTGTGAA 325
Db |
Qy 121 ATTACATCACAAACACCAACATACCTAGCTTCAACACACAGAGTCAAGTCAAACTG 180
Db |
Qy 326 ATTACATCACAAACACCAACATACCTAGCTTCAACACACAGAGTCAAGTCAAACTG 385
Db |
Qy 181 CAACCCCAACAGTCAAGACTTAAACACACAAACACCAACCCCAACACCAACCCCAAGCCC 240
Db |
Qy 386 CAACCCCAACAGTCAAGACTTAAACACACAAACACCAACCCCAACACCAACCCCAAGCCC 445
Qy 241 ACTCAAAAGACGCGCAAAACAAACCCCAACCAACCAACCAATATGATTTTCACTTGAA 300
Db |
Qy 446 ACTCAAAAGACGCGCAAAACAAACCCCAACCAACCAACCAATATGATTTTCACTTGAA 505
Qy 301 GTGTTTAACTTTGTACCTGAGCATATGACAGCAACATCCAACTCTGTGGCTATCTGC 360
Db |
Qy 506 GTGTTTAACTTTGTACCTGAGCATATGACAGCAACATCCAACTCTGTGGCTATCTGC 565
Qy 361 AAAAGAAATACCAAAACAAACAAACCCAGGAAGAAACCAACCAACCCCAACCAACCAACCA 420
Db |
Qy 566 AAAAGAAATACCAAAACAAACAAACCCAGGAAGAAACCAACCAACCCCAACCAACCAACCA 625
Qy 421 ACCTTCAAGACACCAAAAGATCTGAACCTCAACCACTAAACCAAGGAAGTACCC 480
Db |
Qy 626 ACCTTCAAGACACCAAAAGATCTGAACCTCAACCACTAAACCAAGGAAGTACCC 685
Qy 481 ACCCAAGGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCA 540
Db |
Qy 686 ACCCAAGGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCA 745
Qy 541 CTGCTCACAACACACACAGGAATCCAAACCTCAAGTCAAAATGGAACCTTCCAC 600
Db |
```

```
Db 746 CTGCTCACAACACACACAGGAATCCAAACCTCAAGTCAATGGAACCTTCCAC 805
Qy 601 TCAACTCTCTCCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGCACCA 660
Db |
Db 806 TCAACTCTCTCCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGCACCA 865
Qy 661 TCACAAACCTCATCTCCACCCCAACACAGCCAGCTAGTATTAAAAA 715
Db |
Db 866 TCACAAACCTCATCTCCACCCCAACACAGCCAGCTAGTATTAAAAA 920

RESULT 8
US-08-467-963C-28
; Sequence 28, Application US/08467963C
; Patent No. 5968776
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EMASYSVN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,963C
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/838,189
; FILING DATE: 16-APR-1997
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-467-963C-28

Query Match 97.3%; Score 696; DB 2; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.8e-180;
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CACAAAGTCACACTAACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60
Db |
Db 199 CACAAAGTCACACTAACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 258
Qy 61 ACCCAACATACCTCCTCAGATCCTCAGTTGGAATCAGTTCTTCAATCTGTGAA 120
Db |
Db 259 ACCCAACATACCTCCTCAGATCCTCAGTTGGAATCAGTTCTTCAATCTGTGAA 318
```



121 ATTACATCAAAACCAACCACTACTAGTCTTCAACACACAGGAGTCAAGTCAAACTG 180  
122 |||||  
319 ATTACATCAAAACCAACCACTACTAGTCTTCAACACACAGGAGTCAAGTCAAACTG 378  
320 |||||  
181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCA 240  
182 |||||  
379 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCA 438  
380 |||||  
241 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
242 |||||  
439 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498  
440 |||||  
301 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCA 360  
302 |||||  
499 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCA 558  
500 |||||  
361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
362 |||||  
559 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618  
560 |||||  
421 ACCTTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
422 |||||  
619 ACCTTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678  
620 |||||  
481 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
482 |||||  
679 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 738  
680 |||||  
541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
542 |||||  
739 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798  
740 |||||  
601 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 660  
602 |||||  
799 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 858  
800 |||||  
661 TCACACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 696  
662 |||||  
859 TCACACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 894  
858 |||||

## RESULT 9

US-08-838-189D-28  
Sequence 28, Application US/08838189D  
Patent No. 5998169  
GENERAL INFORMATION:  
APPLICANT: KLEIN, Michel H  
APPLICANT: DU, Run-Pan  
APPLICANT: EWASZYHN, Mary E  
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST  
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
City: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1K7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,189D  
FILING DATE: 16-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jfb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-838-189D-28

Query Match 97.3%; Score 696; DB 2; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.8e-180;  
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAAGTCACACTAACTGCAATCATACAGATGCAACAGCCAGATCAAGAAACA 60  
DB 199 CACAAAGTCACACTAACTGCAATCATACAGATGCAACAGCCAGATCAAGAAACA 258  
QY 61 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 120  
DB 259 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 318  
QY 121 ATTACATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180  
DB 319 ATTACATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 378  
QY 181 CAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
DB 379 CAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 438  
QY 241 ACTACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 439 ACTACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498  
QY 301 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCA 360  
DB 499 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCA 558  
QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
DB 559 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618  
QY 421 ACCTTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
DB 619 ACCTTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678  
QY 481 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
DB 679 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 738  
QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
DB 739 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798  
QY 601 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 660  
DB 799 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 858  
QY 661 TCACACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 696  
DB 859 TCACACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 894

## RESULT 10

US-08-852-344D-28

```

1 / Sequence 28, Application US/08852344D
2 / Patent No. 6017539
3 / GENERAL INFORMATION:
4 / APPLICANT: KLEIN, Michel H
5 / APPLICANT: DU, Run-Pan
6 / APPLICANT: EWASYSHYN, Mary E
7 / TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
8 / TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
9 / TITLE OF INVENTION: SYNCYTIAL VIRUS
10 / NUMBER OF SEQUENCES: 37
11 / CORRESPONDENCE ADDRESS:
12 / ADDRESS: Sim & McBurney
13 / STREET: 6th Floor, 330 University Avenue
14 / CITY: Toronto
15 / STATE: Ontario
16 / COUNTRY: Canada
17 / ZIP: M5G 1R7
18 / COMPUTER READABLE FORM:
19 / MEDIUM TYPE: Floppy disk
20 / COMPUTER: IBM PC compatible
21 / OPERATING SYSTEM: PC-DOS/MS-DOS
22 / SOFTWARE: PatentIn Release #1.0, Version #1.30
23 / CURRENT APPLICATION DATA:
24 / APPLICATION NUMBER: US/08/852,344D
25 / FILING DATE: 07-MAY-1997
26 / CLASSIFICATION: 424
27 / PRIOR APPLICATION DATA:
28 / APPLICATION NUMBER: US 08/344,639
29 / FILING DATE: 14-NOV-1994
30 / PRIOR APPLICATION DATA:
31 / APPLICATION NUMBER: GB 9200117.1
32 / FILING DATE: 06-JAN-1992
33 / ATTORNEY/AGENT INFORMATION:
34 / NAME: STEWART, Michael I
35 / REGISTRATION NUMBER: 24,973
36 / REFERENCE/DOCKET NUMBER: 1038-688 MIS:jfb
37 / TELECOMMUNICATION INFORMATION:
38 / TELEPHONE: (416) 595-1155
39 / TELEFAX: (416) 595-1163
40 / INFORMATION FOR SEQ ID NO: 28:
41 / SEQUENCE CHARACTERISTICS:
42 / LENGTH: 894 base pairs
43 / TYPE: nucleic acid
44 / STRANDEDNESS: single
45 / TOPOLOGY: linear
46 /
47 /
48 /
49 /
50 /
51 /
52 /
53 /
54 /
55 /
56 /
57 /
58 /
59 /
60 /
61 /
62 /
63 /
64 /
65 /
66 /
67 /
68 /
69 /
70 /
71 /
72 /
73 /
74 /
75 /
76 /
77 /
78 /
79 /
80 /
81 /
82 /
83 /
84 /
85 /
86 /
87 /
88 /
89 /
90 /
91 /
92 /
93 /
94 /
95 /
96 /
97 /
98 /
99 /
100 /
101 /
102 /
103 /
104 /
105 /
106 /
107 /
108 /
109 /
110 /
111 /
112 /
113 /
114 /
115 /
116 /
117 /
118 /
119 /
120 /
121 /
122 /
123 /
124 /
125 /
126 /
127 /
128 /
129 /
130 /
131 /
132 /
133 /
134 /
135 /
136 /
137 /
138 /
139 /
140 /
141 /
142 /
143 /
144 /
145 /
146 /
147 /
148 /
149 /
150 /
151 /
152 /
153 /
154 /
155 /
156 /
157 /
158 /
159 /
160 /
161 /
162 /
163 /
164 /
165 /
166 /
167 /
168 /
169 /
170 /
171 /
172 /
173 /
174 /
175 /
176 /
177 /
178 /
179 /
180 /
181 /
182 /
183 /
184 /
185 /
186 /
187 /
188 /
189 /
190 /
191 /
192 /
193 /
194 /
195 /
196 /
197 /
198 /
199 /
200 /
201 /
202 /
203 /
204 /
205 /
206 /
207 /
208 /
209 /
210 /
211 /
212 /
213 /
214 /
215 /
216 /
217 /
218 /
219 /
220 /
221 /
222 /
223 /
224 /
225 /
226 /
227 /
228 /
229 /
230 /
231 /
232 /
233 /
234 /
235 /
236 /
237 /
238 /
239 /
240 /
241 /
242 /
243 /
244 /
245 /
246 /
247 /
248 /
249 /
250 /
251 /
252 /
253 /
254 /
255 /
256 /
257 /
258 /
259 /
260 /
261 /
262 /
263 /
264 /
265 /
266 /
267 /
268 /
269 /
270 /
271 /
272 /
273 /
274 /
275 /
276 /
277 /
278 /
279 /
280 /
281 /
282 /
283 /
284 /
285 /
286 /
287 /
288 /
289 /
290 /
291 /
292 /
293 /
294 /
295 /
296 /
297 /
298 /
299 /
300 /
301 /
302 /
303 /
304 /
305 /
306 /
307 /
308 /
309 /
310 /
311 /
312 /
313 /
314 /
315 /
316 /
317 /
318 /
319 /
320 /
321 /
322 /
323 /
324 /
325 /
326 /
327 /
328 /
329 /
330 /
331 /
332 /
333 /
334 /
335 /
336 /
337 /
338 /
339 /
340 /
341 /
342 /
343 /
344 /
345 /
346 /
347 /
348 /
349 /
350 /
351 /
352 /
353 /
354 /
355 /
356 /
357 /
358 /
359 /
360 /
361 /
362 /
363 /
364 /
365 /
366 /
367 /
368 /
369 /
370 /
371 /
372 /
373 /
374 /
375 /
376 /
377 /
378 /
379 /
380 /
381 /
382 /
383 /
384 /
385 /
386 /
387 /
388 /
389 /
390 /
391 /
392 /
393 /
394 /
395 /
396 /
397 /
398 /
399 /
400 /
401 /
402 /
403 /
404 /
405 /
406 /
407 /
408 /
409 /
410 /
411 /
412 /
413 /
414 /
415 /
416 /
417 /
418 /
419 /
420 /
421 /
422 /
423 /
424 /
425 /
426 /
427 /
428 /
429 /
430 /
431 /
432 /
433 /
434 /
435 /
436 /
437 /
438 /
439 /
440 /
441 /
442 /
443 /
444 /
445 /
446 /
447 /
448 /
449 /
450 /
451 /
452 /
453 /
454 /
455 /
456 /
457 /
458 /
459 /
460 /
461 /
462 /
463 /
464 /
465 /
466 /
467 /
468 /
469 /
470 /
471 /
472 /
473 /
474 /
475 /
476 /
477 /
478 /
479 /
480 /
481 /
482 /
483 /
484 /
485 /
486 /
487 /
488 /
489 /
490 /
491 /
492 /
493 /
494 /
495 /
496 /
497 /
498 /
499 /
500 /
501 /
502 /
503 /
504 /
505 /
506 /
507 /
508 /
509 /
510 /
511 /
512 /
513 /
514 /
515 /
516 /
517 /
518 /
519 /
520 /
521 /
522 /
523 /
524 /
525 /
526 /
527 /
528 /
529 /
530 /
531 /
532 /
533 /
534 /
535 /
536 /
537 /
538 /
539 /
540 /
541 /
542 /
543 /
544 /
545 /
546 /
547 /
548 /
549 /
550 /
551 /
552 /
553 /
554 /
555 /
556 /
557 /
558 /
559 /
560 /
561 /
562 /
563 /
564 /
565 /
566 /
567 /
568 /
569 /
570 /
571 /
572 /
573 /
574 /
575 /
576 /
577 /
578 /
579 /
580 /
581 /
582 /
583 /
584 /
585 /
586 /
587 /
588 /
589 /
590 /
591 /
592 /
593 /
594 /
595 /
596 /
597 /
598 /
599 /
600 /
601 /
602 /
603 /
604 /
605 /
606 /
607 /
608 /
609 /
610 /
611 /
612 /
613 /
614 /
615 /
616 /
617 /
618 /
619 /
620 /
621 /
622 /
623 /
624 /
625 /
626 /
627 /
628 /
629 /
630 /
631 /
632 /
633 /
634 /
635 /
636 /
637 /
638 /
639 /
640 /
641 /
642 /
643 /
644 /
645 /
646 /
647 /
648 /
649 /
650 /
651 /
652 /
653 /
654 /
655 /
656 /
657 /
658 /
659 /
660 /
661 /
662 /
663 /
664 /
665 /
666 /
667 /
668 /
669 /
670 /
671 /
672 /
673 /
674 /
675 /
676 /
677 /
678 /
679 /
680 /
681 /
682 /
683 /
684 /
685 /
686 /
687 /
688 /
689 /
690 /
691 /
692 /
693 /
694 /
695 /
696 /
697 /
698 /
699 /
700 /
701 /
702 /
703 /
704 /
705 /
706 /
707 /
708 /
709 /
710 /
711 /
712 /
713 /
714 /
715 /
716 /
717 /
718 /
719 /
720 /
721 /
722 /
723 /
724 /
725 /
726 /
727 /
728 /
729 /
730 /
731 /
732 /
733 /
734 /
735 /
736 /
737 /
738 /
739 /
740 /
741 /
742 /
743 /
744 /
745 /
746 /
747 /
748 /
749 /
750 /
751 /
752 /
```

|                       |                 |                    |           |             |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match           | 97.3%;          | Score 696;         | DB 3;     | Length 894; |
| Best Local Similarity | 100.0%;         | Prd. No. 2.8e-180; |           |             |
| Matches 696;          | Conservative 0; | Mismatches 0;      | Indels 0; | Gaps 0;     |

  

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | CACAAAGTCACACTAACACTGCGAATCATATCAAGATGCAACAAGCCAGATCAAGAACACA  | 60  |
|    |     |  |     |
| Db | 199 | CACAAAGTCACACTAAACAATGCAATCATATCAAGATGCAACAAGCCAGATCAAGAACACA  | 258 |
| QY | 61  | ACCCCAACATACCTTCACCTCAGGATCCCTCAGCTTGGAAATCAGCTTCTCCAATCTCTGAA | 120 |
|    |     |  |     |
| Db | 259 | ACCCCAACATACCTTCACCTCAGGATCCCTCAGCTTGGAAATCAGCTTCTCCAATCTCTGAA | 318 |
| QY | 121 | ATTACATCAAAAACCCACCATACTAGCTTTCACAAACACACGAGATCAAGTCAAAACCTG   | 180 |
|    |     |  |     |
| Db | 319 | ATTACATCAAAAACCCACCATACTAGCTTTCACAAACACACGAGATCAAGTCAAAACCTG   | 378 |
| QY | 181 | CAAGCCCAACAGCTCAAGACTTAAAAACACAAACAACCCCAAAACACAACCCCAAGGCC    | 240 |
|    |     |  |     |
| Db | 379 | CAAGCCCAACAGCTCAAGACTTAAAAACACAAACAACCCCAAAACACAACCCCAAGGCC    | 438 |
| QY | 241 | ACTACAAAAACAAGCCCAAAAACAAACCCACCAAAACCCCAATATGATTTTTCATCTCGAA  | 300 |
|    |     |  |     |
| Db | 439 | ACTACAAAAACAAGCCCAAAAACAAACCCACCAAAACCCCAATATGATTTTTCATCTCGAA  | 498 |
| QY | 301 | GTGTTTAACTTTGTACCCCTGCAGCATATGCAGCAACAATCAACCTGCTGGGCTATCTGC   | 360 |

|     |   |     |
|-----|---|-----|
| 499 | GTGTTTAACCTTTGTATGCGCGCTATGCGAGCAACAATCCAAACCTGCTGTGGGCTATCTGCG   | 555 |
| 361 | AAAGAATACCAAAACAAAAACAGAGAAAGAAAAACCAACCAACCAAGGCTTACAAAAAACCA    | 420 |
| 559 | AAAGAATACCAAAACAAAAACAGAGAAAGAAAAACCAACCAACCAAGGCTTACAAAAAACCA    | 618 |
| 421 | ACCTTCAAGACACGAAAAAAGATCTCAAACTCTAAACGACTTAAACCAAGGAAGTACCC       | 480 |
| 619 | ACCTTCAAGACACGAAAAAAGATCTCAAACTCTAAACGACTTAAACCAAGGAAGTACCC       | 678 |
| 481 | ACCACCAAGCCACAGAGAGGCGCAACCATCAACACCAACCAACCAACCAACCAACCAACCA     | 540 |
| 679 | ACCACCAAGCCACAGAGAGGCGCAACCATCAACACCAACCAACCAACCAACCAACCAACCA     | 738 |
| 541 | CTGCTCACCACCAACCAACCAACCAAGGAAATCCAAAACTCACAAGTCAAAATGGAACCTTCCAC | 600 |
| 739 | CTGCTCACCACCAACCAACCAAGGAAATCCAAAACTCACAAGTCAAAATGGAACCTTCCAC     | 798 |
| 601 | TCACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCCACCAACCTCCGAGCAACCA     | 660 |
| 799 | TCACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCCACCAACCTCCGAGCAACCA     | 858 |
| 661 | TCACCAACCTCATCTCCACCCCAACCAACCAACGCGAG                            | 696 |
| 859 | TCACCAACCTCATCTCCACCCCAACCAACCAACGCGAG                            | 894 |

RESULT 11  
 US-08-344-639E-28  
 ; Sequence 28, Application US/08344639E  
 ; Patent No. 6033668  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Klein, Michel H  
 ; APPLICANT: Du, Run-Pan  
 ; APPLICANT: Ewasysbyn, Mary E  
 ; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS  
 ; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
 ; TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS  
 ; NUMBER OF SEQUENCES: 38  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sim & McBurney  
 ; STREET: 330 University Avenue, 6th Floor  
 ; CITY: Toronto  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: M5G 1R7  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/344,639E  
 ; FILING DATE: 14-NOV-1994  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/001,554  
 ; FILING DATE: 06-JAN-1993  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9200117.1  
 ; FILING DATE: 06-JAN-1992  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Stewart, Michael I  
 ; REGISTRATION NUMBER: 24,973  
 ; REFERENCE/DOCKET NUMBER: 1038-391 MIS  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (416) 595-1155  
 ; TELEFAX: (416) 595-1163  
 ; TELEX: 065-24567 SIMBAS  
 ; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:

LENGTH: 894 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-344-639E-28

Query Match 97.3%; Score 696; DB 3; Length 894;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-180;  
 Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACATTAACAACTGCAATCATCAAGATGCAACAGCCAGATCAAGACACA 60  
 Db CACAAAGTCACATTAACAACTGCAATCATCAAGATGCAACAGCCAGATCAAGACACA 258  
 QY 61 ACCCAACATACCTCACTCAGGATCCTCAGCTTGGAAATCAGTTCTCCAAATCTGTGAA 120  
 Db ACCCAACATACCTCACTCAGGATCCTCAGCTTGGAAATCAGTTCTCCAAATCTGTGAA 318  
 QY 121 ATTACATCAAAACCAACACCATATAGTCTTCAACACACAGAGTCAAGTCAAACTG 180  
 Db ATTACATCAAAACCAACACCATATAGTCTTCAACACACAGAGTCAAGTCAAACTG 378  
 QY 181 CAACCCCAACAGTCAAGATCTTAAACACACACACACACACACACACACACACACAC 240  
 Db CAACCCCAACAGTCAAGATCTTAAACACACACACACACACACACACACACACACAC 438  
 QY 241 ACTACAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 Db ACTACAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498  
 QY 301 GTGTTAACTTTGACCTGAGCATATGAGCAACCAATCCAACTGCTGGCTATCTGC 360  
 Db GTGTTAACTTTGACCTGAGCATATGAGCAACCAATCCAACTGCTGGCTATCTGC 558  
 QY 361 AAAAGATATCAACAAACCAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 420  
 Db AAAAGATATCAACAAACCAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 618  
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480  
 Db ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 678  
 QY 481 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 540  
 Db ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 738  
 QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600  
 Db CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 798  
 QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGCCCA 660  
 Db TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGCCCA 858  
 QY 661 TCACACCTCTATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 696  
 Db TCACACCTCTATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 894

## RESULT 12

US-08-801-898A-23  
 Sequence 23, Application US/08801898A  
 Patent No. 5998602  
 GENERAL INFORMATION:  
 APPLICANT: Torrence, Paul F.  
 APPLICANT: Silverman, Robert H.  
 APPLICANT: Cirino, Nick M.  
 APPLICANT: Li, Guiying  
 APPLICANT: Xiao, Wei  
 TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE  
 TITLE OF INVENTION: OLIGONUCLEOTIDES EFFECTIVE TO TREAT RSV INFECTIONS  
 NUMBER OF SEQUENCES: 29  
 CORRESPONDENCE ADDRESS:

ADDRESS: Pennie & Edmonds LLP  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: NY  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/801,898A  
 FILING DATE: 18-FEB-1997  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Poissant, Brian M.  
 REGISTRATION NUMBER: 28,462  
 REFERENCE/DOCKET NUMBER: 8656-009  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 23:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 15222 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 FEATURE:  
 NAME/KEY: RSV-A2  
 LOCATION: 1...15222  
 OTHER INFORMATION:  
 US-08-801-898A-23

Query Match 91.6%; Score 654.8; DB 2; Length 15222;  
 Best Local Similarity 94.8%; Pred. No. 1.5e-168;  
 Matches 677; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 CACAAAGTCACATTAACAACTGCAATCATCAAGATGCAACAGCCAGATCAAGACACA 60  
 Db CACAAAGTCACATTAACAACTGCAATCATCAAGATGCAACAGCCAGATCAAGACACA 4945  
 QY 61 ACCCAACATACCTCACTCAGGATCCTCAGTTGGAAATCAGTTCTCCAAATCTGTGAA 120  
 Db ACCCAACATACCTCACTCAGGATCCTCAGTTGGAAATCAGTTCTCCAAATCTGTGAA 5005  
 QY 121 ATTACATCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180  
 Db ATTACATCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5065  
 QY 181 CAACCCCAACAGTCAAGATCTAAGATCAACCAACCAACCAACCAACCAACCAACCAAC 240  
 Db CAATCCCAACAGTCAAGATCTAAGATCAACCAACCAACCAACCAACCAACCAACCAAC 5125  
 QY 241 ACTACAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300  
 Db ACCCAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5185  
 QY 301 GTGTTTAACTTTGACCTGAGCATATGCAACCAACCAACCAACCAACCAACCAACCAAC 360  
 Db GTGTTTAACTTTGACCTGAGCATATGCAACCAACCAACCAACCAACCAACCAACCAAC 5245  
 QY 361 AAAAGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 Db AAAAGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5305  
 QY 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
 Db ACCCTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5365  
 QY 481 ACCACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540



```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-892-403A-1

Query Match          91.6%; Score 654.8; DB 2; Length 15223;
Best Local Similarity 94.8%; Pred. No. 1.5e-168;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTGCAATCATACAGATGCAACCAAGCCGATCAAGAACACA 60
DB 4887 CACAAGTCACCAACCACTGCAATCATACAGATGCAACCAAGCCGATCAAGAACACA 4946

QY 61 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAAATCAGTCTTCCAAATCTCTGAA 120
DB 4947 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAAATCAGTCTTCCAAATCTCTGAA 5006

QY 121 ATTACATCAACCAACCACTCACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 180
DB 5007 ATTACATCAACCAACCACTCACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 5066

QY 181 CAACCCCAACAGTCAGATCTTAAACACACACACCAACCAACCAACCAACCAACCA 240
DB 5067 CAATCCCAACAGTCAGATCTTAAACACACACCAACCAACCAACCAACCAACCA 5126

QY 241 ACTCAAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 5127 ACCAACAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186

QY 301 GTGTTTAACTTTGATCTGAGCATATGAGCAACCAATCAACCTCTGGCTGATCTGC 360
DB 5187 GTGTTTAACTTTGATCTGAGCATATGAGCAACCAATCAACCTCTGGCTGATCTGC 5246

QY 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 5247 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5306

QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 5307 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5366

QY 481 ACCACCAAGCCCAAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCAACCA 540
DB 5367 ACCACCAAGCCCAAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCAACCA 5426

QY 541 CTGCTCACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 5427 CTGCTCACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5486

QY 601 TCAACCTCTCTCCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGCAACCA 660
DB 5487 TCAACCTCTCTCCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGTACCCA 5546

QY 661 TCACACCTCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 714
DB 5547 TCACACCTCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5600

RESULT 15
US-08-720-132-1
; Sequence 1, Application US/08720132
; Patent No. 6264957
; GENERAL INFORMATION:
; APPLICANT: Collins, Peter L.
; TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY
; TITLE OF INVENTION: SYNCTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
```

```
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,132
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/007,083
; FILING DATE: 27-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 15280-250-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15223 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other
US-08-720-132-1
```

```
Query Match          91.6%; Score 654.8; DB 3; Length 15223;
Best Local Similarity 94.8%; Pred. No. 1.5e-168;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTGCAATCATACAGATGCAACCAAGCCGATCAAGAACACA 60
DB 4887 CACAAGTCACACCAACCACTGCAATCATACAGATGCAACCAAGCCGATCAAGAACACA 4946

QY 61 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAAATCAGTCTTCCAAATCTCTGAA 120
DB 4947 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAAATCAGTCTTCCAAATCTCTGAA 5006

QY 121 ATTACATCAACCAACCACTCACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 180
DB 5007 ATTACATCAACCAACCACTCACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTG 5066

QY 181 CAACCCCAACAGTCAGATCTTAAACACACACCAACCAACCAACCAACCAACCA 240
DB 5067 CAATCCCAACAGTCAGATCTTAAACACACCAACCAACCAACCAACCAACCA 5126

QY 241 ACTCAAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 5127 ACCAACAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186

QY 301 GTGTTTAACTTTGATCTGAGCATATGAGCAACCAATCAACCTCTGGCTGATCTGC 360
DB 5187 GTGTTTAACTTTGATCTGAGCATATGAGCAACCAATCAACCTCTGGCTGATCTGC 5246

QY 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 5247 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5306

QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 5307 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5366

QY 481 ACCACCAAGCCCAAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCAACCA 540
DB 5367 ACCACCAAGCCCAAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCAACCA 5426

QY 541 CTGCTCACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 5427 CTGCTCACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5486
```

|    |      |   |      |
|----|------|---|------|
| Qy | 601  | TCAACTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACATCCGAGCAGCCA | 660  |
|    |      |   |      |
| Db | 5487 | TCAACTTCTCCGAGGCAATCCAGCCCTTCTCAAGTCTTACAAATCCGAGTACCCA   | 5546 |
|    |      |   |      |
| Qy | 661  | TCACAACCTCATCTCCACCCCAACAACACAGCCAGTAGTTATTAAAAA          | 714  |
|    |      |   |      |
| Db | 5547 | TCACAACCTTCATCTCCACCCCAACAACACAGCCAGTAGTTACTTAAAAACATA    | 5600 |
|    |      |   |      |

Search completed: October 30, 2003, 01:17:38  
Job time : 57.5382 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 16:50:52 ; Search time 244.018 Seconds  
(without alignments)  
7909.644 Million cell updates/sec

Title: US-09-462-816-3

Perfect score: 715

Sequence: 1 cacaagaagtcactaacac.....gtagttataaaaaaaaaa 715

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03:\*

|     |   |
|-----|---|
| 1:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*  |
| 2:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*  |
| 3:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*  |
| 4:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*  |
| 5:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*  |
| 6:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*  |
| 7:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*  |
| 8:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*  |
| 9:  | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*  |
| 10: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*  |
| 11: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*  |
| 12: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*  |
| 13: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*  |
| 14: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*  |
| 15: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*  |
| 16: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*  |
| 17: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*  |
| 18: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*  |
| 19: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*  |
| 20: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*  |
| 21: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*  |
| 22: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.* |
| 23: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.* |
| 24: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*  |
| 25: | /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 715   | 100.0       | 715    | 20 | AXX08422    |
| 2          | 715   | 100.0       | 920    | 14 | AAQ45686    |
| 3          | 715   | 100.0       | 920    | 20 | AXX08421    |
| 4          | 659   | 92.2        | 935    | 8  | AAV70784    |
| 5          | 659   | 92.2        | 935    | 19 | AAV18736    |
| 6          | 657.4 | 91.9        | 935    | 13 | AAQ29623    |
| 7          | 654.8 | 91.6        | 15210  | 20 | AXX59703    |
| 8          | 654.8 | 91.6        | 15222  | 18 | AAV78440    |

|    |       |      |       |    |          |                      |
|----|-------|------|-------|----|----------|----------------------|
| 9  | 654.8 | 91.6 | 15223 | 18 | AAV63430 | Respiratory syncyt   |
| 10 | 654.8 | 91.6 | 15223 | 19 | AAV17553 | Respiratory syncyt   |
| 11 | 654.8 | 91.6 | 15223 | 21 | AAV88743 | Respiratory syncyt   |
| 12 | 649.6 | 90.9 | 918   | 22 | AAV88494 | Human RSV G-protei   |
| 13 | 646.2 | 90.4 | 897   | 24 | AAV20145 | Respiratory syncyt   |
| 14 | 445.8 | 62.3 | 696   | 19 | AAV38298 | Respiratory syncyt   |
| 15 | 300.4 | 42.0 | 15225 | 19 | AAV17552 | Respiratory syncyt   |
| 16 | 300.4 | 42.0 | 15225 | 21 | AAV88744 | Respiratory syncyt   |
| 17 | 300.4 | 42.0 | 15225 | 24 | AAV25520 | Human respiratory    |
| 18 | 295.6 | 41.3 | 15229 | 19 | AAV18276 | RSV isolate 18537    |
| 19 | 295.6 | 41.3 | 15229 | 20 | AAV22910 | Nucleotide sequenc   |
| 20 | 295.6 | 41.3 | 15229 | 20 | AAV35268 | DNA encoding the L   |
| 21 | 273.6 | 38.3 | 15218 | 19 | AAV18275 | RSV isolate 2B wil   |
| 22 | 273.6 | 38.3 | 15218 | 20 | AAV22909 | Nucleotide sequenc   |
| 23 | 273.6 | 38.3 | 15218 | 20 | AAV35267 | Nucleotide sequenc   |
| 24 | 273.6 | 38.3 | 15218 | 20 | AAV35267 | DNA encoding the L   |
| 25 | 273.6 | 38.3 | 15219 | 19 | AAV18277 | RSV vaccine 2B33F    |
| 26 | 273.6 | 38.3 | 15219 | 19 | AAV18278 | RSV vaccine 2B20L    |
| 27 | 273.6 | 38.3 | 15219 | 19 | AAV18279 | RSV revertant 2B33   |
| 28 | 273.6 | 38.3 | 15219 | 19 | AAV18280 | RSV revertant 2B20   |
| 29 | 273.6 | 38.3 | 15219 | 20 | AAV22911 | Nucleotide sequenc   |
| 30 | 273.6 | 38.3 | 15219 | 20 | AAV22912 | Nucleotide sequenc   |
| 31 | 273.6 | 38.3 | 15219 | 20 | AAV22913 | Nucleotide sequenc   |
| 32 | 273.6 | 38.3 | 15219 | 20 | AAV35269 | DNA encoding the L   |
| 33 | 273.6 | 38.3 | 15219 | 20 | AAV35270 | DNA encoding the L   |
| 34 | 273.6 | 38.3 | 15219 | 20 | AAV35271 | DNA encoding the L   |
| 35 | 273.6 | 38.3 | 15219 | 20 | AAV35272 | DNA encoding the L   |
| 36 | 212.8 | 29.8 | 1050  | 17 | AAV31647 | Encodes Streptococ   |
| 37 | 212.8 | 29.8 | 1050  | 22 | AAV84711 | Nucleotide sequenc   |
| 38 | 212.8 | 29.8 | 1356  | 22 | AAV80153 | Nucleotide sequenc   |
| 39 | 210.8 | 29.5 | 303   | 16 | AAV03486 | RSV subgroup A clo   |
| 40 | 210.8 | 29.5 | 303   | 17 | AAV27066 | RSV subgroup A prote |
| 41 | 210.8 | 29.5 | 303   | 17 | AAV27073 | RSV subgroup A wi    |
| 42 | 210.8 | 29.5 | 303   | 17 | AAV31649 | Respiratory Syncyt   |
| 43 | 210.8 | 29.5 | 303   | 20 | AAV30478 | RSV G protein anti   |
| 44 | 210.8 | 29.5 | 303   | 21 | AAV55882 | DNA encoding a G2N   |
| 45 | 210.8 | 29.5 | 303   | 22 | AAV78459 | Nucleotide sequenc   |

#### ALIGNMENTS

RESULT 1  
ID AAX08422 standard; cDNA; 715 BP.  
XX  
AC AAX08422;  
XX  
DT 28-JUN-1999 (first entry)  
XX  
DE G protein gene fragment of respiratory syncytial virus.  
XX  
KW G protein; respiratory syncytial virus; RSV; recombinant vector;  
KW vaccine; immune response; immunogenicity; tPA; antibody;  
KW tissue plasminogen activator; ss.  
XX  
OS Respiratory syncytial virus (RSV).  
XX  
FH Key Location/Qualifiers  
FT CDS 1..702  
FT /tag= a  
FT /product= "Secreted G protein"  
XX  
PN WO9904010-A1.  
XX  
PD 28-JAN-1999.  
XX  
PF 16-JUL-1998; 98WO-CA00697.  
XX  
PR 18-JUL-1997; 97US-0896442.  
XX  
PA (CONN-) CONNAUGHT LAB LTD.  
XX



PI Klein MH, Li X, Sambhara S;  
XX WPI; 1999-132254/11.  
DR P-PSDB; AAW96314.  
XX Immunogenic composition for generating antibodies against  
PT respiratory syncytial virus - comprises non-replicating vector  
PT containing the protein G sequence, useful in protective vaccines and  
PT to raise antibodies for diagnosis  
XX  
XX Claim 8; Figure 3; 67pp; English.  
XX  
CC The respiratory syncytial virus (RSV) G protein can be used in  
CC vaccines by inserting the G protein gene into a non-replicating  
CC vector. The G protein is placed under the control of alternative  
CC signal and expression sequences, for example the chimeric G protein  
CC produced may also comprise the signal peptide of tissue plasminogen  
CC activator (tPA). The recombinant vector may also comprise sequences  
CC upstream of the G protein gene which enhance the G proteins  
CC immunoprotective ability. The resulting immunogenic composition will  
CC generate antibodies directed against the RSV G protein when  
CC administered to a host organism. The composition is useful as a  
CC vaccine to immunise against RSV-associated disease, particularly  
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
CC by usual immunisation and cell fusion methods.  
XX  
XX Sequence 715 BP; 305 A; 240 C; 70 G; 100 T; 0 other;  
SQ  
Query Match 100.0%; Score 715; DB 20; Length 715;  
Best Local Similarity 100.0%; Pred. No. 1.4e-162;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAAGTCACACTAACCTGCAATCATCATACAGATGCAACAGCCAGATCAAGACACA 60  
Db 1 CACAAAGTCACACTAACCTGCAATCATCATACAGATGCAACAGCCAGATCAAGACACA 60  
QY 61 ACCCCAAATACCTCACTCAGGATCTCTGGAATCAGCTTCTCCAATCTGTCTGAA 120  
Db 61 ACCCCAAATACCTCACTCAGGATCTCTGGAATCAGCTTCTCCAATCTGTCTGAA 120  
QY 121 ATTACATCAAAACCCACCACTACTAGTTCACACACACAGAGTCAAGTCAAACTG 180  
Db 121 ATTACATCAAAACCCACCACTACTAGTTCACACACACAGAGTCAAGTCAAACTG 180  
QY 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240  
Db 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240  
QY 241 ACTCAAAACCAACGCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 241 ACTCAAAACCAACGCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
QY 301 GTGTTTAACTTTCTACCTGACATATGACGATCAGCAATCCACCTGCTGGGTATCTGC 360  
Db 301 GTGTTTAACTTTCTACCTGACATATGACGATCAGCAATCCACCTGCTGGGTATCTGC 360  
QY 361 AAAAGATATCAAAACCAAAACCAAGGAAACCAACCAACCAACCAACCAACCAACCA 420  
Db 361 AAAAGATATCAAAACCAAAACCAAGGAAACCAACCAACCAACCAACCAACCAACCA 420  
QY 421 ACCTTCAAGCAACCAACCAAAAGATCTCAAAACCTCAAAACCAACCAACCAAGGATACCC 480  
Db 421 ACCTTCAAGCAACCAACCAAAAGATCTCAAAACCTCAAAACCAACCAAGGATACCC 480  
QY 481 ACCACCAAGCCACAGAGAGCCACCATCAACACCAACCAACCAACCAACCAACCAACCA 540  
Db 481 ACCACCAAGCCACAGAGAGCCACCATCAACACCAACCAACCAACCAACCAACCAACCA 540  
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
Db 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
QY 601 TCAACCTCTCTCGAAGGCAATCTTAAGCCCTTCTCAAGTCTTCCAAACATCCGAGACCCA 660

Db 601 TCAACCTCTCTCGAAGGCAATCTTAAGCCCTTCTCAAGTCTTCCAAACATCCGAGACCCA 660  
QY 661 TCACAAACCTCTCTCTCCACCCACACACACACGCGCAGTAGTATTAAAAA 715  
Db 661 TCACAAACCTCTCTCTCCACCCACACACACACGCGCAGTAGTATTAAAAA 715  
RESULT 2  
AAQ45686  
ID AAQ45686 standard; DNA; 920 BP.  
XX  
XX AAQ45686;  
XX AC  
XX 25-MAR-2003 (updated)  
DT 13-JAN-1994 (first entry)  
XX  
XX Respiratory syncytial virus (RSV) G gene.  
XX PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine;  
XX db.  
XX Respiratory syncytial virus.  
XX OS  
XX Key Location/Qualifiers  
XX CDS 8..901  
XX /tag= a  
XX /product= RSV G protein  
XX misc\_feature 89..176  
XX /tag= b  
XX /label= Transmembrane anchor domain  
XX  
XX WO9314207-AL.  
XX  
XX 22-JUL-1993.  
XX  
XX 05-JAN-1993; 93WO-CA00001.  
XX  
XX 06-JAN-1992; 92GB-0000117.  
XX  
XX (CONN-) CONNAUGHT LAB LTD.  
XX  
XX Ewasysbyn ME, Klein MH;  
XX  
XX WPI; 1993-243222/30.  
XX P-PSDB; AAR39286.  
XX  
XX Multimeric hybrid genes and their chimeric proteins - are  
XX PT vaccines against multiple pathogenic infections e.g.  
XX PT para-influenza virus and respiratory syncytial virus  
XX  
XX Claim 11; Figure 7A-7D; 80pp; English.  
XX  
XX A novel multimeric hybrid gene is used as a vaccine. The gene  
XX consists of two gene sequences which are linked and encode antigenic  
XX regions, these two sequences being derived from two different  
XX pathogens (para-influenza virus (PIV) and respiratory syncytial virus  
XX (RSV)). The gene sequences that are particularly used are those  
XX which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
XX G proteins (AAQ45685, AAQ45686).  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;  
SQ  
Query Match 100.0%; Score 715; DB 14; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.5e-162;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAAGTCACACTAACCTGCAATCATCATACAGATGCAACAGCCAGATCAAGACACA 60  
Db 206 CACAAAGTCACACTAACCTGCAATCATCATACAGATGCAACAGCCAGATCAAGACACA 265  
QY 61 ACCCCAAATACCTCACTCAGGATCTCTGGAATCAGCTTCTCCAATCTGTCTGAA 120



QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 660  
 DB 806 TCAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 865  
 QY 661 TCACAAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 715  
 DB 866 TCACAAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 920

RESULT 4  
 AAN70784  
 ID AAN70784 standard; cDNA; 935 BP.  
 XX  
 AC AAN70784;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 05-APR-1991 (first entry)  
 XX  
 XX Sequence encoding human respiratory syncytial virus (HRSV) A2 strain  
 DE G protein.  
 DE  
 XX  
 XX Vaccine; ss.  
 XX  
 XX Human respiratory syncytial virus (HRSV).  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 16..912  
 FT /\*tag= a  
 FT  
 XX  
 PN WO8704185-A.  
 XX  
 PD 16-JUL-1987.  
 XX  
 PF 23-DEC-1986; 86WO-US02756.  
 XX  
 PR 14-JAN-1986; 86US-0818740.  
 XX  
 PA (UNYNC-) UNIV NORTH CAROLINA.  
 PA (WERTZ/) WERTZ G W.  
 XX  
 DR WPI; 1987-206300/29.  
 DR P-PSDB; AAP70845.  
 XX  
 PT Vaccines for human respiratory virus - comprising proteins or  
 PT fragment encoded by a DNA sequence coding for human respiratory  
 PT syncytial virus proteins.  
 XX  
 PS Disclosure; Chart 13; 57pp; English.  
 XX  
 CC A novel plasmid which comprises a DNA sequence encoding this  
 CC protein, and the protein itself, are claimed, for use as HRSV  
 CC vaccines. The vaccine can be administered to pregnant women or to  
 CC women of child bearing age to stimulate maternal antibodies.  
 CC Infants can also be vaccinated at 2-3 months of age.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match 92.2%; Score 659; DB 8; Length 935;  
 Best Local Similarity 95.1%; Pred. No. 4.5e-149;  
 Matches 680; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAATCGAATCATACAGATGCAACAGCCAGATCAAGAACA 60  
 DB 214 CACAAAGTCACACTAATCGAATCATACAGATGCAACAGCCAGATCAAGAACA 273  
 QY 61 ACCGCAACACT 120  
 DB 274 ACCGCAACACT 333  
 QY 121 ATTACATCAACACACACACATCTAGCTTCAACACACAGGATCAAGTCAACCTG 180

DB 334 ATTACATCAACATCACCACCATCTAGCTTCAACACACAGGATCAAGTCAACCTG 393  
 QY 181 CAAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240  
 DB 394 CAAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 453  
 QY 241 ACTACAAACAGGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 300  
 DB 454 ACCACAAACAGGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 513  
 QY 301 GTGTTTAACTTTGTATCCCTGAGCATATGAGCAACATCCCAACCTGCTGGGTATCTGC 360  
 DB 514 GTGTTCAACTTTGTATCCCTGAGCATATGAGCAACATCCCAACCTGCTGGGTATCTGC 573  
 QY 361 AAAAGAAATACCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 420  
 DB 574 AAAAGAAATACCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 633  
 QY 421 ACCTTCAAGACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480  
 DB 634 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693  
 QY 481 ACCACCAAGCCACAGAGAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
 DB 694 ACCACCAAGCCACAGAGAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 753  
 QY 541 CTGCTCACAAC 600  
 DB 754 CTACTCACCTCCAC 813  
 QY 601 TCACACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 660  
 DB 814 TCACACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 873  
 QY 661 TCACACCT 715  
 DB 874 TCACACCT 928

RESULT 5  
 AAV18736  
 ID AAV18736 standard; cDNA; 935 BP.  
 XX  
 AC AAV18736;  
 XX  
 DT 11-JUN-1998 (first entry)  
 XX  
 DE HRSV glycoprotein G cDNA.  
 XX  
 KW HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine; ss.  
 XX  
 OS Human respiratory syncytial virus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 16..912  
 FT /\*tag= a  
 FT  
 FT /product= glycoprotein\_G  
 XX  
 PN US5716823-A.  
 XX  
 PD 10-FEB-1998.  
 XX  
 PF 12-MAY-1997; 97US-0854783.  
 XX  
 PR 13-JUL-1988; 88US-0218737.  
 PR 14-JAN-1986; 86US-0818740.  
 PR 23-DEC-1986; 86WO-US02756.  
 PR 11-JUN-1992; 92US-0897171.  
 PR 12-MAY-1997; 97US-0854783.  
 XX  
 PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Collins PL, Wertz GW;

```

XX WP1; 1998-144802/13.
DR P-PSDB; AAW47605.
XX
XX Production of human respiratory syncytial virus glyco-protein F or G
PT -by culturing eukaryotic host cells transfected with corresponding
PT DNA
XX
XX Example 1; Columns 27-28; 17pp; English.
XX
XX The present sequence was used in the development of a novel method
CC for the production of human respiratory syncytial virus (HRSV)
CC glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises
CC culturing eukaryotic host cells transfected with an isolated DNA
CC sequence encoding HRSV gpF or gpG. The gp can be used to prepare
CC vaccines against HRSV.
XX
XX Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;
SQ
Query Match 92.2%; Score 659; DB 19; Length 935;
Best Local Similarity 95.1%; Pred. No. 4.5e-149;
Matches 680; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 1 CACAAAGTCACTAACTAACTCAATCATATAGATGCAAGCCAGATCAAGAACACA 60
DB 214 CACAAAGTCACTAACTAACTCAATCATATAGATGCAAGCCAGATCAAGAACACA 273
QY 61 ACCCAACATPACTCTCACTCAGGATCCTCAGTTGGAAATCAGTTCTCCAAATCTGTCTGAA 120
DB 274 ACCCAACATPACTCTCACTCAGGATCCTCAGTTGGAAATCAGTTCTCCAAATCTGTCTGAA 333
QY 121 ATTATCATCAATACCACTCACTAGTCTTCAATCAACCACTCAAGTCAAGTCAAGCTG 180
DB 334 ATTATCATCAATACCACTCACTAGTCTTCAATCAACCACTCAAGTCAAGTCAAGCTG 393
QY 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
DB 394 CAATCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 453
QY 241 ACTACAAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 454 ACCCAACAAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513
QY 301 GTGTTTAACTTGTACCTCTCAGCATATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 360
DB 514 GTGTTTAACTTGTACCTCTCAGCATATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 573
QY 361 AAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 574 AAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
DB 634 ACCCTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 693
QY 481 ACCCAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
DB 694 ACCCAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 753
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
DB 754 CTACTCACTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 813
QY 601 TCAACCTCTCTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 660
DB 814 TCAACCTCTCTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 873
QY 661 TCACCAACCTCTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB 874 TCACCAACCTCTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 928

```

RESULT 6

```

AAQ29623
ID AAQ29623 standard; DNA; 935 BP.
XX
XX AAQ29623;
AC
XX 25-MAR-2003 (updated)
DT 03-MAR-1993 (first entry)
DT
XX
XX HSRV glycoprotein G (gpG).
XX
XX Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
XX major capsid protein; N; ss.
XX
XX Human respiratory syncytial virus strain A2.
XX
XX Key Location/Qualifiers
FT CDS 16..912
FT FT /*tag= a
FT FT /label= G_protein
FT FT misc_feature 16..22
FT FT /*tag= b
FT FT /note= "Oligonucleotide used to probe for
FT FT full length cDNA"
FT FT misc_feature 898..912
FT FT /*tag= c
FT FT /note= "Oligonucleotide used to specifically
FT FT prime the reverse transcription reaction
FT FT for making the first strand of the cDNA"
FT
PN US5149650-A.
XX
XX 22-SEP-1992.
PD
XX
XX 13-JUL-1988; 88US-0218737.
PF
XX
XX 14-JAN-1986; 86US-0818740.
PR
XX 13-JUL-1988; 88US-0218737.
XX
XX (UYN-C) UNIV NORTH CAROLINA.
PA
XX
XX Collins PL, Wertz GW;
PI
XX
XX WP1; 1992-340247/41.
DR P-PSDB; AAR25302.
XX
XX Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
PT fragments
PT
PS Disclosure; Page 18; 21pp; English.
XX
XX The sequences of mRNA encoding HRSV structural proteins are given in
CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid
CC protein N. The sequences and encoded proteins are useful for
CC preparing vaccines against HRSV. The vaccines can be used to confer
CC immunity against respiratory tract infections on human subjects.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 935 BP; 382 A; 294 C; 100 G; 159 T; 0 other;
SQ
Query Match 91.9%; Score 557.4; DB 13; Length 935;
Best Local Similarity 95.0%; Pred. No. 1.1e-148;
Matches 679; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1 CACAAAGTCACTAACTAACTCAATCATATAGATGCAAGCCAGATCAAGAACACA 60
DB 214 CACAAAGTCACTAACTAACTCAATCATATAGATGCAAGCCAGATCAAGAACACA 273
QY 61 ACCCAACATPACTCTCACTCAGGATCCTCAGTTGGAAATCAGTTCTCCAAATCTGTCTGAA 120
DB 274 ACCCAACATPACTCTCACTCAGGATCCTCAGTTGGAAATCAGTTCTCCAAATCTGTCTGAA 333
QY 121 ATTATCATCAATACCACTCACTAGTCTTCAATCAACCACTCAAGTCAAGTCAAGCTG 180

```

```

Db 334 ATTACATCAAAATCACCACTACTAGCTTCAACCAACAGGCGTCAAGTCAACCCCTG 393
QY 181 CAAACCAACAGTCAAGCTAAAGCTAAACACAAACAAACCAACCCAGAGGCC 240
Db 394 CAATCCACAAAGTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 453
QY 241 ACTACAAACAGCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 454 ACCACAAACAGCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513
QY 301 GTGTTTAACTTTGTACCTCGAGCATATGCAAGCAACAAATCCAACTGCTGGCTATCTGC 360
Db 514 GTGTTCACTTTGTACCTCGAGCATATGCAAGCAACAAATCCAACTGCTGGCTATCTGC 573
QY 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
Db 574 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
Db 634 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 693
QY 481 ACCACCAAGCCCAACAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCA 540
Db 694 ACCACCAAGCCCAACAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCA 753
QY 541 CTGCTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
Db 754 CTACTCACCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 813
QY 601 TCAACTCTCTCGAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGCCCA 660
Db 814 TCAACTCTCTCGAGGCAATCTAAGCCCTTCTCAAGTCTTACACACATCCGAGTACCA 873
QY 661 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 715
Db 874 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 928

```

## RESULT 7

AAK59703/C  
ID AAK59703 standard; DNA; 15210 BP.

XX  
AC AAK59703;

DT 26-JUL-1999 (first entry)

XX  
XX Polynucleotide sequence of RSV strain A2.

DE  
XX Antisense oligonucleotide; negative-strand RNA virus; activator; RNase L;  
KW respiratory syncytial virus; RSV; influenza; mumps; rabies; ss.

XX  
XX Respiratory syncytial virus.

XX  
XX WO9922742-A1.

XX  
XX 14-MAY-1999.

XX  
XX 02-NOV-1998; 98WO-US23391.

XX  
XX 03-NOV-1997; 97US-0962690.

XX  
XX (CLEV-) CLEVELAND CLINIC FOUND.

PA (USSH ) US NAT INST OF HEALTH.

XX  
XX Cirino NM, Li G, Player MR, Silverman RH, Torrence PF;

PI Xiao W;

XX  
XX WPI; 1999-326917/27.

XX  
XX New composition useful for inhibiting or treating infections against  
PT negative-strand RNA virus

XX  
PS Disclosure; Fig 1; 98pp; English.

XX  
CC The specification describes a composition comprising a polynucleotide  
CC consisting of an antisense oligonucleotide containing a hydroxy group,  
CC complementary to the genomic or antigenomic strand of a negative-strand  
CC RNA virus; and an activator of RNase L. The polynucleotide is used to  
CC inhibit, or treat, infection by negative-strand RNA viruses, specifically  
CC respiratory syncytial virus (RSV) but also (para)influenza, mumps, and  
CC rabies. The polynucleotide can cross cell membranes without requiring  
CC carriers or permeabilizing agents, and can selectively cleave the RNA  
CC targeted by the oligonucleotide. The present sequence represents  
CC the polynucleotide sequence of RSV strain A2.

SQ Sequence 15210 BP; 4232 A; 2351 C; 2700 G; 5919 T; 8 other;

Query Match 91.6%; Score 654.8; DB 20; Length 15210;

Best Local Similarity 94.8%; Pred. No. 9.4e-148;

Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

QY 1 CACAAAGTCACTAACCACTGCAATCATACAAGATGCAACAGCCAGATCAGAACACA 60
Db 10331 CACAAAGTCACTAACCACTGCAATCATACAAGATGCAACAGCCAGATCAGAACACA 10272
QY 61 ACCCAACATACCTCACTCAGGATCCTCAGCTTGAATCAGCTTCCAAATCTGTGAA 120
Db 10271 ACCCAACATACCTCACTCAGGATCCTCAGCTTGAATCAGCTTCCAAATCTGTGAA 10212
QY 121 ATTACATCAAAACCAACCAACCAATAGCTTCAACCAACAGAGGTCAAGTCAACCTG 180
Db 10211 ATTACATCAAAATCAACCAACCAATAGCTTCAACCAACAGAGGTCAAGTCAACCTG 10152
QY 181 CAAACCAACAGTCAAGTCAAGCTTAAACCAACCAACCAACCAACCAACCAACCAAC 240
Db 10151 CAAACCAACAGTCAAGTCAAGCTTAAACCAACCAACCAACCAACCAACCAACCAAC 10092
QY 241 ACTCAAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
Db 10091 ACTCAAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 10032
QY 301 GTGTTTAACTTTGTACCTCGAGCATATGCAAGCAACCAATCCAACTGCTGGCTATCTGC 360
Db 10031 GTGTTTAACTTTGTACCTCGAGCATATGCAAGCAACCAATCCAACTGCTGGCTATCTGC 9972
QY 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 420
Db 9971 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 9912
QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
Db 9911 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 9852
QY 481 ACCCAAGGCCCAACAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCAACCA 540
Db 9851 ACCCAAGGCCCAACAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCAACCA 9792
QY 541 CTGCTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
Db 9791 CTGCTCACCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 9732
QY 601 TCAACTCTCTCGAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGCCCA 660
Db 9731 TCAACTCTCTCGAGGCAATCTAAGCCCTTCTCAAGTCTTCCACACATCCGAGTACCA 9672
QY 661 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 714
Db 9671 TCACAACTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 9618

```

## RESULT 8

AAK78440

ID AAK78440 standard; DNA; 15222 BP.

XX

AC AAT78440;  
 XX 17-MAR-1998 (first entry)  
 XX Human respiratory syncytial virus strain A2.  
 XX Antisense oligonucleotide; respiratory syncytial virus;  
 KW RSV; treatment; infection; inhibition; strain A2; ss.  
 XX Human respiratory syncytial virus.  
 OS WO9729757-A1.  
 XX 21-AUG-1997.  
 XX 14-FEB-1997; 97WO-US02531.  
 XX 15-FEB-1996; 96US-0011725.  
 XX (CLEV-) CLEVELAND CLINIC FOUND.  
 PA (USSH ) US NAT INST OF HEALTH.  
 XX Cirino NM, Li G, Silverman RH, Torrence PF, Xiao W;  
 XX WPI; 1997-424748/39.  
 DR Polynucleotide containing sequence anti-sense to region of RSV -  
 XX connected via a linker to an activator of RNaseL, used to treat RSV  
 PT infections  
 XX Disclosure; Pages 47-51; 89pp; English.  
 XX The present sequence was used in the preparation of a novel  
 CC polynucleotide, comprising an antisense oligonucleotide, with a  
 CC hydroxy group at one end, that is complementary to 15-20 bases of  
 CC the anti-genomic RNA strand of a respiratory syncytial virus (RSV),  
 CC a linker attached to the OH-end of the antisense oligonucleotide and  
 CC an oligonucleotide activator of RNaseL attached to the linker. The  
 CC polynucleotide can be used to treat RSV infections, which can also  
 CC be treated by administration of the antisense oligonucleotide, so as  
 CC to form a complex with activated RNase L in vivo. The  
 CC polynucleotide can be transported across the cell membranes without  
 CC carriers or permeability agents, and once introduced destroys  
 CC antisense target RNA. It also inhibits RSV infection in vitro in a  
 CC superior manner to the conventional drug, ribavirin.  
 XX Sequence 15222 BP; 5923 A; 2706 C; 2356 G; 4237 T; 0 other;  
 SQ  
 Query Match 91.6%; Score 654.8; DB 18; Length 15222;  
 Best Local Similarity 94.8%; Pred. No. 9.4e-148;  
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
 QY 1 CACAAAGTCACACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60  
 Db 4886 CACAAAGTCACACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 4945  
 QY 61 ACCCCACATACCTCAGTCAGGATCCCTCAGCTTGGAATCAGCTTCTCAATCTGCTGAA 120  
 Db 4946 ACCCCACATACCTCAGTCAGGATCCCTCAGCTTGGAATCAGCTTCTCAATCTGCTGAA 5005  
 QY 121 ATTACATCACAACACCCACCACCTACTAGTCTTCAACACACAGGATCAAGTCAAACTG 180  
 Db 5006 ATTACATCACAACATCACCACCTACTAGTCTTCAACACACAGGATCAAGTCAAACTG 5065  
 QY 181 CAACCCACACAGTCAAGATCAAAACACACACACACACACACACACACACACACACACCC 240  
 Db 5066 CAATCCACACAGTCAAGATCAAAACACACACACACACACACACACACACACACACCC 5125  
 QY 241 ACTACAAACACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 Db 5126 ACCACAAACACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5185  
 QY 301 GTGTGTTAACTTTGACCTCGAGCATATGCGACGACCAACCAACCTGCTGGCTATCTGC 360

Db 5186 GTGTTCACTTTGTACCTCGAGCATATGCGACGACCAACCTGCTGGCTATCTGC 5245  
 QY 361 AAAAGAATACAAACAAAAACCCAGGAAAGAAAAACCCACCAAGCCTTACAAAAAACCA 420  
 Db 5246 AAAAGAATACCAACAAAAACCCAGGAAAGAAAAACCCACCAAGCCTTACAAAAAACCA 5305  
 QY 421 ACCTTCAAGACAAACAAAAAGATCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCT 480  
 Db 5306 ACCCTCAAGACAAACAAAAAGATCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCT 5365  
 QY 481 ACCACCAAGCCACAGAGAGCCAAACCATCAACACACACACACACACACACACACACATCA 540  
 Db 5366 ACCACCAAGCCACAGAGAGCCAAACCATCAACACACACACACACACACACACACATCA 5425  
 QY 541 CTGCTCAACCAACACACACAGAGAAATCCAAAACTCAAAAGTCAAAATGGAACCTTCCAC 600  
 Db 5426 CTACTCACTCCAAACACACAGAGAAATCCAGAACTCAAAAGTCAAAATGGAACCTTCCAC 5485  
 QY 601 TCRAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCCAACATCCGAGACCCCA 660  
 Db 5486 TCRAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCTCAACATCCGAGTACCCA 5545  
 QY 661 TCACACCTCTCATCTCCACCCCAACACACACACACACACACACACACACACATCA 714  
 Db 5546 TCACACCTCTCATCTCCACCCCAACACACACACACACACACACACACACATCA 5599  
 RESULT 9  
 AAT63430  
 ID AAT63430 standard; DNA, 15223 BP.  
 XX  
 AC AAT63430;  
 XX  
 DT 02-JUL-1997 (first entry)  
 XX Respiratory syncytial virus anti-genome.  
 XX RSV; vaccine; gene therapy; upper respiratory tract disease; ss.  
 XX Human respiratory syncytial virus strain A2.  
 XX WO9712032-A1.  
 XX 03-APR-1997.  
 XX 27-SEP-1996; 96WO-US15524.  
 XX 27-SEP-1995; 95US-0007083.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Collins PL;  
 XX WPI; 1997-212893/19.  
 XX Infectious respiratory syncytial virus particles - useful for  
 PT treatment of RSV or gene therapy of upper respiratory tract diseases  
 XX Claim 46; Page 43-51; 66pp; English.  
 XX A human respiratory syncytial virus (RSV) anti-genome sequence  
 CC (AAT63430) is the 5' to 3' positive-sense sequence of RSV; the genome  
 CC itself is negative-sense. It was synthesised in segments by RT-PCR  
 CC using intracellular RSV mRNA or genomic rRNA isolated from purified  
 CC viruses as template. Restriction site markers were introduced by  
 CC incorporating the changes into the primers used for RT-PCR. The  
 CC recombinant sequence can be expressed with a nucleocapsid protein,  
 CC a nucleocapsid phosphoprotein, a large polymerase protein and an  
 CC RNA elongation factor to produce isolated infectious RSV particles  
 CC useful for generating vaccines against RSV. Recombinant RSV genome  
 CC or antigenome can also be used as a vector for gene therapy of the  
 CC upper respiratory tract.

XX SQ Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;  
 Query Match 91.6%; Score 654.8; DB 18; Length 15223;  
 Best Local Similarity 94.8%; Pred. No. 9.4e-148;  
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTCACTGCAATCATACAGATGCAACAGCCGAGTCAAGAACACA 60  
 DB CACAAAGTCACACTCACTGCAATCATACAGATGCAACAGCCGAGTCAAGAACACA 4946  
 QY 61 ACCCAACATACCTCACTCAGATCTCTCAGTTCGAAATCAGTCTCTCAATCTCTGAA 120  
 DB ACCCAACATACCTCACTCAGATCTCTCAGTTCGAAATCAGTCTCTCAATCTCTGAA 5006  
 QY 121 ATTACATCAAAACACCAACATCTAGCTTCAACACACAGGAGTCAAGTCAAACTG 180  
 DB ATTACATCAAAATCACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTG 5066  
 QY 181 CAACCCACACAGTCAGATCAATCAAAACACACACACCAACCAACCAACCAAGCC 240  
 DB CAATCCACACAGTCAGATCAATCAAAACACACACACCAACCAACCAACCAAGCC 5126  
 QY 241 ACTCAAAACACACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 DB ACCAACAACACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186  
 QY 301 GTGTTTAACTTTGACCTGAGCATATGCAAGCAACCAATCCACTGCTGGCTATCTGC 360  
 DB GTGTTTAACTTTGACCTGAGCATATGCAAGCAACCAATCCACTGCTGGCTATCTGC 5246  
 QY 361 AARAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 DB AARAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5306  
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
 DB ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5366  
 QY 481 ACCACCAAGCCACAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
 DB ACCACCAAGCCACAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCA 5426  
 QY 541 CTGCTCAGCAACACACACAGGAAATCAAAATCTCAAGTCAAAATGGAACCTTCCAC 600  
 DB CTGCTCAGCAACACACAGGAAATCAAAATCTCAAGTCAAAATGGAACCTTCCAC 5486  
 QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCGAGCACCCA 660  
 DB TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCGAGTACCCA 5546  
 QY 661 TCACACCTCTATCTCCACCAACACACAGGAGCCAGTGTATTATTAACCAACCA 714  
 DB TCACACCTCTATCTCCACCAACACACAGGAGCCAGTGTATTATTAACCAACCA 5546

RESULT 10  
 AAV17553  
 ID AAV17553 standard; cDNA; 15223 BP.  
 AC AAV17553;  
 XX AAV17553;  
 XX AAV17553;  
 DT 20-JUL-1998 (first entry)  
 DE Respiratory syncytial virus antigenome.  
 KW RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.  
 XX Human respiratory syncytial virus D46.  
 OS WO9802530-A1.  
 FN 22-JAN-1998.  
 PD

XX PF 15-JUL-1997; 97WO-US12269.  
 XX PR 23-MAY-1997; 97US-0047634.  
 PR 15-JUL-1996; 96US-0021773.  
 PR 09-MAY-1997; 97US-0046141.  
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 PA Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN,  
 PI Whitehead SS;  
 XX WPI; 1998-110579/10.  
 DR Attenuated respiratory syncytial virus vaccines - useful to protect  
 PT individuals against RSV infection  
 PT Example 7; Page 188-195; 238pp; English.  
 XX This is the 5'-3' positive sequence nucleotide sequence of  
 CC respiratory syncytial virus (RSV) D46. The genome is  
 CC negative-sense; the complete nucleotide sequence of the  
 CC wild-type B-1 virus has also been determined (see AAV17552).  
 CC A novel infectious recombinant RSV comprises a RSV genome or  
 CC antigenome, a major nucleocapsid (N) protein, a nucleocapsid  
 CC phosphoprotein (P), a large polymerase protein (L), and a RNA  
 CC polymerase elongation factor, where the recombinant RSV has at  
 CC least two attenuating mutations, one of the mutations specifying a  
 CC temperature-sensitive (ts) substitution at amino acid Phe521,  
 CC Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts  
 CC nucleotide substitution in the gene-start sequence of gene M2.  
 CC Also claimed are: (i) an isolated infectious RSV particle which  
 CC comprises a recombinant RSV (anti)genome, N, P, and L proteins, a  
 CC RNA polymerase elongation factor, where the (anti)genome is modified:  
 CC (i) to ablate or modulate expression of a SH, NS1, NS2 or G gene or  
 CC a cis-acting regulatory sequence; and (ii) by a termination codon  
 CC introduced within a selected gene, or by a change in sequence,  
 CC position or presence of a GS or GE transcription signal relative to  
 CC the selected gene; (2) an expression vector; and (3) an RSV strain  
 CC selected from cpts RSV 248 (ATCC VR 2450), cpts RSV 530 (ATCC VR  
 CC 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR  
 CC 2454), cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or  
 CC B-1 cp52/285 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated  
 CC attenuated recombinant RSV and RSV particles are used in a vaccine  
 CC to stimulate the immune system of an individual to induce  
 CC protection against RSV. The expression vector of (2) is used for  
 CC the production of infectious attenuated RSV particles.  
 XX SQ Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;  
 Query Match 91.6%; Score 654.8; DB 19; Length 15223;  
 Best Local Similarity 94.8%; Pred. No. 9.4e-148;  
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTCACTGCAATCATACAGATGCAACAGCCGAGTCAAGAACACA 60  
 DB CACAAAGTCACACTCACTGCAATCATACAGATGCAACAGCCGAGTCAAGAACACA 4946  
 QY 61 ACCCAACATACCTCACTCAGATCTCTCAGTTCGAAATCAGTCTCTCAATCTCTGAA 120  
 DB ACCCAACATACCTCACTCAGATCTCTCAGTTCGAAATCAGTCTCTCAATCTCTGAA 5006  
 QY 121 ATTACATCAAAACACCAACATCTAGCTTCAACACACAGGAGTCAAGTCAAACTG 180  
 DB ATTACATCAAAATCACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTG 5066  
 QY 181 CAACCCACACAGTCAGATCAATCAAAACACACACACCAACCAACCAACCAAGCC 240  
 DB CAATCCACACAGTCAGATCAATCAAAACACACACACCAACCAACCAACCAAGCC 5126  
 QY 241 ACTCAAAACACACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 DB ACCAACAACACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186  
 QY 301 GTGTTTAACTTTGACCTGAGCATATGCAAGCAACCAATCCACTGCTGGCTATCTGC 360  
 DB GTGTTTAACTTTGACCTGAGCATATGCAAGCAACCAATCCACTGCTGGCTATCTGC 5246  
 QY 361 AARAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 DB AARAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5306  
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
 DB ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5366  
 QY 481 ACCACCAAGCCACAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
 DB ACCACCAAGCCACAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCA 5426  
 QY 541 CTGCTCAGCAACACACAGGAAATCAAAATCTCAAGTCAAAATGGAACCTTCCAC 600  
 DB CTGCTCAGCAACACACAGGAAATCAAAATCTCAAGTCAAAATGGAACCTTCCAC 5486  
 QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCGAGCACCCA 660  
 DB TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCGAGTACCCA 5546  
 QY 661 TCACACCTCTATCTCCACCAACACACAGGAGCCAGTGTATTATTAACCAACCA 714  
 DB TCACACCTCTATCTCCACCAACACACAGGAGCCAGTGTATTATTAACCAACCA 5546



```

QY 301 GTGTTTAACTTTGTACCTGAGCATATGAGCAAAATCAACCTGCTGGGCTATCTGC 360
Db 5187 GTGTTCAACTTTGTACCTGAGCATATGAGCAAAATCAACCTGCTGGGCTATCTGC 5246
QY 361 AAAAGATACCAACCAAAACCAAGGAAAGAAAGAAACCAACCAAGGCTACAAAACCA 420
Db 5247 AAAAGATACCAACCAAAACCAAGGAAAGAAAGAAACCAACCAAGGCTACAAAACCA 5306
QY 421 ACCTTCAAGCAACCAAAAGAGTCTCAAACTCAAAACCACTAAACCAAGGAGTACCC 480
Db 5307 ACCTTCAAGCAACCAAAAGAGTCTCAAACTCAAAACCACTAAACCAAGGAGTACCC 5366
QY 481 ACCACCAAGCCCAAGAGAGCCCAACCATCAACCAACCAACCAACCAACCACTACA 540
Db 5367 ACCACCAAGCCCAAGAGAGCCCAACCATCAACCAACCAACCAACCAACCACTACA 5426
QY 541 CTGCTACCAACCAACCAAGGAAATCCAAATCTCAAGTCAAAATGGAACCTTCCAC 600
Db 5427 CTACTACCTCCCAACCAACCAAGGAAATCCAAATCTCAAGTCAAAATGGAACCTTCCAC 5486
QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCCAACCATCCGAGCACCCA 660
Db 5487 TCAACTTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCCAACCATCCGAGTACCCA 5546
QY 661 TCACACCTCATCTCCACCAACCAACCAAGGCGAGTGTATTAAAAAAA 714
Db 5547 TCACACCTCATCTCCACCAACCAACCAAGGCGAGTGTATTAAAAACATA 5600

RESULT 11
AAA88743
ID AAA88743 standard; cDNA; 15223 BP.
AC
XX
AC AAA88743;
XX
DT 19-FEB-2001 (first entry)
XX
DE Respiratory syncytial virus D46 5'-3' positive sense sequence.
XX
KW RSV; vaccine; attenuation; pneumonia; bronchiolitis; mutant; ss.
XX
OS Chimeric - Human respiratory syncytial virus.
OS Chimeric - Bacteriophage T7.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT variation replace(4,G)
FT mutation 1099
FT /*tag= a
FT /*tag= b
FT /*note= "single C insertion to create AfIII site
FT mutation replace(1139,A)
FT /*tag= c
FT /*note= "creates NcoI site in N gene nontranslated
FT mutation replace(1140,G)
FT /*tag= d
FT /*note= "creates NcoI site in N gene nontranslated
FT mutation replace(5612,A)
FT /*tag= e
FT /*note= "creates StuI site in G/F intergenic region"
FT mutation replace(5616,A)
FT /*tag= f
FT /*note= "creates StuI site in G/F intergenic region"
FT mutation replace(7560,A)
FT /*tag= g
FT /*note= "creates SphI site in P/M2 intergenic region"
XX
PN WO200061611-A2.
XX

```

```

PD 19-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08802.
XX
PR 13-APR-1999; 99US-0291894.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Collins PL, Murphy BR, Whitehead SS;
XX
DR WPI; 2000-679462/66.
XX
FT Infectious chimeric respiratory syncytial virus (RSV) produced from
FT cloned nucleotide sequences, useful as a vaccine against diseases
FT caused by the virus, such as pneumonia and bronchiolitis -
XX
PS Example 7; Page 262-268; 280pp; English.
XX
CC The present sequence is that of the 5' to 3' positive-sense
CC sequence of human respiratory syncytial virus (RSV); the genome
CC itself is negative-sense. This antigenome cDNA, termed D46, was
CC synthesized in segments by RT-PCR using synthetic oligonucleotides
CC as primers and intracellular RSV mRNA or genome RNA isolated from
CC purified virions as template. The antigenome includes a 5'-terminal
CC nonviral G triplet contributed by the T7 promoter, 4 sequence
CC markers (see AAA88745-47) at positions 1099 (which adds 1 nucleotide
CC to the length), 1139, 5611 and 7559, a ribozyme and tandem T7
CC terminators, and a single nonviral 3'-phosphorylated residue
CC contributed to the 3' end by ribozyme cleavage. The invention
CC provides an isolated infectious chimeric RSV comprising a major
CC nucleocapsid (N) protein, a nucleocapsid phosphoprotein (P), a
CC large polymerase protein (L), an RNA polymerase elongation factor,
CC and a partial or complete RSV genome or antigenome of one RSV
CC strain or subgroup virus combined with a heterologous gene of a
CC different RSV strain or subgroup virus. The chimeric RSV is
CC infectious and attenuated, preferably by introduction of selected
CC mutations. It is useful as a vaccine against RSV, which causes
CC diseases such as pneumonia and bronchiolitis in infants. The
CC immune system of an individual is stimulated to induce protection
CC against natural RSV infection, preferably in a multivalent manner
CC to achieve protection against multiple RSV strains and/or subgroups.
XX
SQ Sequence 15223 BP; 5921 A; 2705 C; 2361 G; 4236 T; 0 other;

Query Match 91.6%; Score 654.8; DB 21; Length 15223;
Best Local Similarity 94.8%; Pred. No. 9.4e-148;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTGCAATCATATCAAGATGCAACCAAGCCAGATCAAGAACCA 60
Db 4887 CACAAAGTCACACCACTGCAATCATATCAAGATGCAACCAAGCCAGATCAAGAACCA 4946
QY 61 ACCCAACATACCTCACTCAGGATCCTCAGTTGGAATCAGTTCTCCAATCTGTTGAA 120
Db 4947 ACCCAACATACCTCAGGATCCTCAGTTGGAATCAGTTCTCCAATCTGTTGAA 5006
QY 121 ATTATCATCAAAACCAACCACTACTAGTTCACAAACCAAGGAGTCAAGTCAACCTG 180
Db 5007 ATTATCATCAAAATCACCACCATATAGTTCACAAACCAAGGAGTCAAGTCAACCTG 5066
QY 181 CAAACCAACCAAGTCAAGACTAAACCAACCAACCAACCAACCAACCAACCAACCA 240
Db 5067 CAATCCAAACAGTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5126
QY 241 ACTACAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5127 ACCACAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186
QY 301 GGTGTTAACTTTGTACCTGAGCATATGAGCAAAATCAACCTGCTGGGCTATCTGC 360
Db 5187 GGTGTTCAACTTTGTACCTGAGCATATGAGCAAAATCAACCTGCTGGGCTATCTGC 5246
QY 361 AAAAGATACCAACCAAAACCAAGGAAAGAAAGAAACCAACCAAGGCTACAAAACCA 420

```

Db 5247 AAAAGAAATACCAACCAAAAAACAGGAAAGAAAAACCTACCAAGCCCAAAAAAACCA 5306  
Qy 421 ACCTTTCAAGACCAACCAAAAAAGATCTCAAACTCAAAACCACTAAACCAAGGAAGTACCC 480  
Db 5307 ACCCTCAAGACCAACCAAAAAAGATCCCAAACTCAAAACCACTAAATCAAGGAAGTACCC 5366  
Qy 481 ACCACCAAGCCCAAGAGAGGCAACCAATCAACCAACCAACCAACCAACCAATCAAACTACA 540  
Db 5367 ACCACCAAGCCCAAGAGAGGCAACCAATCAACCAACCAACCAACCAACCAATCAAACTACA 5426  
Qy 541 CTGCTCACCACCAACCAACCAAGGAATCCAAACTCAAGAGTCAAAATGGAACCTTCCAC 600  
Db 5427 CTACTCACTCCACCAACCAAGGAATCCAAACTCAAGAGTCAAAATGGAACCTTCCAC 5486  
Qy 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGCACCCA 660  
Db 5487 TCAACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCTACCAACATCCGAGTACCCA 5546  
Qy 661 TCACAACTCTATCTCCACCAACCAACCAAGGAGTCTTATTAATAAAAAA 714  
Db 5547 TCACAACTCTATCTCCACCAACCAACCAAGGAGTCTTATTAATAAAAAA 5600

RESULT 12

AAC88494  
ID AAC88494 standard; DNA; 918 BP.  
XX  
AC AAC88494;  
XX  
DT 12-MAR-2001 (first entry)  
XX  
DE Human RSV G-protein gene.  
XX  
KW Chimeric; respiratory syncytial virus; RSV; immunize; ds.  
XX  
OS Unidentified.  
XX  
FN WO200068392-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 10-MAY-2000; 2000WO-US12582.  
XX  
PR 11-MAY-1999; 99US-0133536.  
XX  
PA (UNIT I) UNIV ILLINOIS FOUND.  
XX  
PI Bueto DE, Korban SS, Sandhu J, Krasnyanski SF;  
XX WPI; 2001-122707/13.  
DR  
PT Chimeric nucleic acid construct for immunizing animals and humans  
PT against respiratory syncytial virus (RSV), comprises a sequence adapted  
PT for expression in plants and a RSV protein or peptide coding sequence  
PT  
PS  
PS Disclosure; Fig 10; 67pp; English.  
CC  
CC The present invention relates to a chimeric nucleic acid construct  
CC comprising: a nucleotide sequence adapted for protein expression in  
CC plants; and a respiratory syncytial virus (RSV) coding sequence  
CC encoding an RSV protein or an antigenic protein or peptide of RSV.  
CC The construct can be used to immunize animals and humans against  
CC respiratory syncytial virus. The use of transgenic plants to  
CC generate the antigen allows the production of greater amounts of  
CC antigen.

SQ Sequence 918 BP; 366 A; 293 C; 100 G; 159 T; 0 other;  
Query Match 90.9%; Score 649.6; DB 22; Length 918;  
Best Local Similarity 95.2%; Pred. No. 8.1e-147;  
Matches 670; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1 CACAAAGTCACTAAACAACCTGCATATATACAGATGCAACAAGCCAGATCAAGAACA 60  
Db 214 CACAAAGTCACTAAACAACCTGCATATATACAGATGCAACAAGCCAGATCAAGAACA 273  
Qy 61 ACCCCACATACCTCAGTCCAGGATCCCTCAGCTTGGATTCAGCTTCCATCTCTGAA 120  
Db 274 ACCCCACATACCTCAGTCCAGGATCCCTCAGCTTGGATTCAGCTTCCATCTCTGAA 333  
Qy 121 ATTACATCAAAAACCAACCATCTAGCTTCAACAACCAAGGAGTCAAGTCAAACTG 180  
Db 334 ATTACATCAAAAACCAACCATCTAGCTTCAACAACCAAGGAGTCAAGTCAAACTG 393  
Qy 181 CAACCCACAGAGTCAAGATTAATAACAACAACAACCAACCAACCAAGGAGTCAAGG 240  
Db 394 CAATCCCAACAGTCAAGACCAAAAACCAACAACAACAACCAACCAAGGAGTCAAGG 453  
Qy 241 ACTACAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 454 ACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513  
Qy 301 GTGTTTAACTTTGTACCTCGAGCATATGCAAGCAACAATCCAACTGTGGGCTATCTGC 360  
Db 514 GTGTTTAACTTTGTACCTCGAGCATATGCAAGCAACAATCCAACTGTGGGCTATCTGC 573  
Qy 361 AAAAGATACCAACCAAAAACCAAGGAAAGAAACCAACCAAGGAGTCAAAAACCA 420  
Db 574 AAAAGATACCAACCAAAAACCAAGGAAAGAAACCAACCAAGGAGTCAAAAACCA 633  
Qy 421 ACCTTCAAGACCAACCAAAAACCAAGGAGTCAAAAACCAACCAAGGAGTCAAAAACCA 480  
Db 634 ACCTTCAAGACCAACCAAAAACCAAGGAGTCAAAAACCAACCAAGGAGTCAAAAACCA 693  
Qy 481 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
Db 694 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 753  
Qy 541 CTGCTCACCACCAACCAAGGAAATCCAAACCTCAAGTCAAAATGGAACCACTTCCAC 600  
Db 754 CTGCTCACCACCAACCAAGGAAATCCAAACCTCAAGTCAAAATGGAACCACTTCCAC 813  
Qy 601 TCACCTCTCCGAGGCAATCTAAGCCCTTCAAGTCTCCAGTCTCCAGTCCGAGCACCCA 660  
Db 814 TCACCTCTCCGAGGCAATCTAAGCCCTTCAAGTCTCCAGTCTCCAGTCCGAGCACCCA 873  
Qy 661 TCACACCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 704  
Db 874 TCACACCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 917

RESULT 13

AAS20145  
ID AAS20145 standard; DNA; 897 BP.  
XX  
AC AAS20145;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Respiratory syncytial virus G protein DNA.  
XX  
KW RSV; ds; G protein; heavily glycosylated protein; antianemic; antiviral;  
KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;  
KW antiviral chemotherapeutic compound; humoral response;  
KW cellular immune response; hPIV; paediatric respiratory disease;  
KW globulin gene transfer; sickle cell disease; beta-thalassaemia;  
KW human immunodeficiency virus infection; HIV.  
OS Human respiratory syncytial virus.  
XX  
FH Key Location/Qualifiers  
CDS 1..897  
FT /\*tag= a  
FT /product= "G protein"

XX WO200192548-A2.  
 XX 06-DEC-2001.  
 XX 22-MAY-2001; 2001WO-US16610.  
 XX 01-JUN-2000; 2000US-208701P.  
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX Portner A, Takimoto T;  
 XX WPI; 2002-130534/17.  
 XX P-PSDB; AAU74676.  
 XX Recombinant Sendai virus useful in vaccines to protect infection by  
 XX paramyxoviruses, comprises exogenous nucleic acid encoding  
 XX paramyxovirus protein or its antigenic fragment -  
 XX Disclosure; Page 47; 57pp; English.  
 XX The invention relates to a recombinant Sendai virus comprising an  
 XX exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its  
 XX antigenic fragment. The virus may be administered in combination  
 XX with an antiviral chemotherapeutic compound. Two or more viruses  
 XX expressing different PMV proteins may be co-administered. Compositions  
 XX comprising the virus are useful for eliciting a humoral and/or  
 XX cellular immune response to a PMV in a mammal, particularly a human.  
 XX Further a recombinant Sendai virus comprising an exogenous nucleic acid  
 XX encoding a second PMV protein is also administered and priming and/or  
 XX boosting humoral or cellular immune response comprises administering  
 XX one or more of a recombinant or isolated PMV protein or its antigenic  
 XX fragment, a DNA vaccine encoding the same, and a non-Sendai viral  
 XX vector encoding a PMV protein. The recombinant virus is useful as an  
 XX effective vaccine against hPIV or RSV (the major causes of paediatric  
 XX respiratory disease) and also to express any gene of  
 XX interest in target cells, providing a positive medical impact on  
 XX impaired cells. Wild-type globin gene transfer (i.e. gene therapy)  
 XX into stem cells effects a cure for sickle cell disease or beta-  
 XX thalassaemia. The recombinant virus may also prove effective in  
 XX conferring immunity to human immunodeficiency virus (HIV) infection.  
 XX The Sendai virus replicates at level that is high enough to  
 XX induce sufficient immunity, but does not cause any harm to human  
 XX recipient. The present sequence encodes a respiratory syncytial  
 XX virus (RSV) G protein (heavily glycosylated protein), a PMV protein  
 XX suitable for expression by the recombinant virus of the invention.  
 XX SQ Sequence 897 BP; 359 A; 289 C; 95 G; 154 T; 0 other;  
 Query Match 90.4%; Score 646.2; DB 24; Length 897;  
 Best Local Similarity 95.3%; Pred. No. 5.3e-146;  
 Matches 666; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
 QY 1 CCAAAAGTCACATCAACATCGCAATCATACAGATGCAAGCCGAGATCAAGAACACA 60  
 Db 199 CACAAAGTCACACCAACATCGCAATCATACAGATGCAAGCCGAGATCAAGAACACA 258  
 QY 61 ACCCAACATACCTCACTCAGGATCTCAGCTTGGATCAGCTTCCCAATCTGCTGAA 120  
 Db 259 ACCCAACATACCTCACTCAGGATCTCAGCTTGGATCAGCTTCCCAATCTGCTGAA 318  
 QY 121 ATTACATCAACAAACCAACCACTACTAGCTTCAACAAACCAAGGAGTCAAGTCAACCTG 180  
 Db 319 ATTACATCAACATCAACCACTACTAGCTTCAACAAACCAAGGAGTCAAGTCAACCTG 378  
 QY 181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240  
 Db 379 CAATCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 438  
 QY 241 ACTACAAAACAGCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 Db 439 ACCAACAACAGCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 498

QY 301 GTGTTAACTTGTACCTGCGAGCATATGCGAGCAACATCAACCTGCTGGGCTATCTGC 360  
 Db 499 GTGTTCAACTTTGTACCTGCGAGCATATGCGAGCAACATCAACCTGCTGGGCTATCTGC 558  
 QY 361 AAAAGGAATACCAAAACAAAAACCCAGGAAGAAAAACCCACCAAGCCTTACAAAAAACCA 420  
 Db 559 AAAAGGAATACCAAAACAAAAACCCAGGAAGAAAAACCCACCAAGCCTTACAAAAAACCA 618  
 QY 421 ACCTTCAAGACAAACCAAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 480  
 Db 619 ACCTTCAAGACAAACCAAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 678  
 QY 481 ACCACCAAGCCACAGAGGACCAACCATCAACAGCAGCAACCAACCAACCAACCAACCA 540  
 Db 679 ACCACCAAGCCACAGAGGACCAACCATCAACAGCAGCAACCAACCAACCAACCAACCA 738  
 QY 541 CTGCTCACCACCAACCAACCAAGGAATCCAAATTCACAAAGTCAAAATGGAACCTTCCAC 600  
 Db 739 CTACTCACTCCACACACCAAGGAATCCAAATTCACAAAGTCAAAATGGAACCTTCCAC 798  
 QY 601 TCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGCAACCA 660  
 Db 799 TCAACTTCTCGAAGGCAATCCAGCCCTTCTCAAGTCTCTACAACTCCGAGTACCA 858  
 QY 661 TCACAACTCTCATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 699  
 Db 859 TCACAACTCTCATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 897  
 RESULT 14  
 AAV38298  
 ID AAV38298 standard; DNA; 696 BP.  
 XX AAV38298;  
 AC AAV38298;  
 DT 26-OCT-1998 (first entry)  
 XX Respiratory syncytial virus glycoprotein G gene portion.  
 DE RSV; glycoprotein G; pneumococcal surface protein A; PspA;  
 KW infection; Streptococcus pneumoniae; sepsis; otitis media;  
 KW meningitis; bacteraemia; pneumonia; vaccine; genetic immunisation;  
 KW ss.  
 XX Respiratory syncytial virus.  
 XX W09824927-AL.  
 XX 11-JUN-1998.  
 XX 04-DEC-1997; 97WO-US22847.  
 XX 04-DEC-1996; 96US-0759505.  
 XX (UYAL-) UNIV ALABAMA.  
 XX Briles DE, Curriel DT, McDaniel LS;  
 XX WPI; 1998-333343/29.  
 XX Plasmid containing pneumococcal epitope for expression in eukaryotic  
 XX cells - useful for eliciting immunological response to pneumococcal  
 XX infection or sepsis  
 XX Example 1; Fig 1C; 47pp; English.  
 XX This is a portion of the respiratory syncytial virus glycoprotein G  
 XX (RSVG) gene. It has been inserted into plasmid pCDNA3 (see AAV38297)  
 XX to create plasmid pG74. This plasmid contains a human  
 XX cytomegalovirus immediate early promoter and the RSVG gene portion  
 XX such that when an in-frame fusion is made, the resultant fusion  
 XX protein may be transported to, and anchored in, a mammalian cell

CC membrane where it can be exposed to the host immune system.  
 CC Insertion of pneumococcal surface protein A (PspA) coding sequence  
 CC created plasmid pKS2601. Intramuscular immunisation of BALB/c  
 CC mice with pKS2601 induced protection against an otherwise lethal  
 CC challenge with a capsular type 3 pneumococcus. A claimed plasmid  
 CC for expression of pneumococcal epitope DNA in eukaryotic cells  
 CC includes a promoter for driving expression in a eukaryotic cell  
 CC (e.g., HCMV-IE), DNA encoding a leader sequence (e.g. of RSV)  
 CC and DNA encoding a pneumococcal epitope such as PspA. The invention  
 CC also provides a vaccine comprising the plasmid and a suitable  
 CC carrier or diluent, and optionally one or more cytokines or DNA  
 CC encoding them, or a bacterial delivery system. The vaccine is used  
 CC to elicit an immunological response in a host, including humans,  
 CC susceptible to pneumococcal infection or sepsis. The plasmid can  
 CC also be used to express a pneumococcal epitope of interest in vitro.  
 CC  
 XX SQ Sequence 696 BP; 284 A; 211 C; 78 G; 123 T; 0 other;  
 Query Match 62.3%; Score 445.8; DB 19; Length 696;  
 Best Local Similarity 95.3%; Pred. No. 8.4e-98;  
 Matches 470; Conservative 0; Mismatches 22; Indels 1; Gaps 1;  
 QY 1 CACAAAGTCACACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60  
 Db 199 CACAAAGTCACACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 258  
 QY 61 ACCCAACATACCTCACTCAGAGTCTCAGTTGGAATCAGTTCTTCCAAATCTGTGAA 120  
 Db 259 ACCCAACATACCTCACTCAGAGTCTCAGTTGGAATCAGTTCTTCCAAATCTGTGAA 318  
 QY 121 ATTACATCAAAACCAACCACTACTAGTTTCACACACACAGAGTCAAGTCAAACTTG 180  
 Db 319 ATTACATCAAAATCACCACTACTAGTTTCACACACACAGAGTCAAGTCAAACTTG 378  
 QY 181 CAACCCACACAGTCACAGCTTAAACACACAAAC-AAACCCACACACACCCAGCAAGCC 239  
 Db 379 CAATCCACACAGTCACAGCTTAAACACACAAAC-AAACCCACACACACCCAGCAAGCC 438  
 QY 240 CACTACAAACCAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 299  
 Db 439 CACCACAAACCAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498  
 QY 300 AGTGTCTTAACTTTGACCTGCAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 359  
 Db 499 AGTGTCTTAACTTTGACCTGCAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 558  
 QY 360 CAAAGGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 419  
 Db 559 CAAAGGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618  
 QY 420 AACTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 479  
 Db 619 AACTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678  
 QY 480 CACCACCAAGCCC 492  
 Db 679 CACCACCAAGCCC 691  
 RESULT 15  
 ID AAV17552 standard; cDNA; 15225 BP.  
 XX AAV17552;  
 AC AAV17552;  
 DT 20-JUL-1998 (first entry)  
 XX Respiratory syncytial virus genome.  
 DE RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.  
 XX Human respiratory syncytial virus B-1.  
 OS  
 XX

PN WO9802530-A1.  
 XX 22-JAN-1998.  
 XX 15-JUL-1997; 97WO-US12269.  
 XX 23-MAY-1997; 97US-0047634.  
 PR 15-JUL-1996; 96US-0021773.  
 PR 09-MAY-1997; 97US-0046141.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN;  
 PI Whitehead SS;  
 XX MPI; 1998-110579/10.  
 XX Attenuated respiratory syncytial virus vaccines - useful to protect  
 PT individuals against RSV infection  
 XX  
 PS Example 4; Page 195-202; 238pp; English.  
 XX This is the complete nucleotide sequence of the wild-type B-1  
 CC respiratory syncytial virus (RSV). The genome is negative-sense;  
 CC the 5'-3' positive-sense sequence of D46 is provided in AAV17553.  
 CC A novel infectious recombinant RSV comprises a RSV genome or  
 CC antigenome, a major nucleocapsid (N) protein, a nucleocapsid  
 CC phosphoprotein (P), a large polymerase protein (L), and a RNA  
 CC polymerase elongation factor, where the recombinant RSV has at  
 CC least two attenuating mutations, one of the mutations specifying a  
 CC temperature-sensitive (ts) substitution at amino acid Phe521,  
 CC Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts  
 CC nucleotide substitution in the gene-start sequence of gene M2.  
 CC Also claimed are: (1) an isolated infectious RSV particle which  
 CC comprises a recombinant RSV (antigenome, N, P, and L proteins, a  
 CC RNA polymerase elongation factor, where the (antigenome) is modified:  
 CC (i) to ablate or modulate expression of a SH, NS1, NS2 or G gene or  
 CC a cis-acting regulatory sequence; and (ii) by a termination codon  
 CC introduced within a selected gene; or by a change in sequence,  
 CC position or presence of a GS or GE transcription signal relative to  
 CC the selected gene; (2) an expression vector; and (3) an RSV strain  
 CC selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR  
 CC 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452),  
 CC cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or  
 CC B-1 cp52/2B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated  
 CC attenuated recombinant RSV and RSV particles are used in a vaccine  
 CC to stimulate the immune system of an individual to induce  
 CC protection against RSV. The expression vector of (2) is used for  
 CC the production of infectious attenuated RSV particles.  
 XX  
 SQ Sequence 15225 BP; 5904 A; 2710 C; 2399 G; 4212 T; 0 other;  
 Query Match 42.0%; Score 300.4; DB 19; Length 15225;  
 Best Local Similarity 64.8%; Pred. No. 1.8e-62;  
 Matches 464; Conservative 0; Mismatches 251; Indels 3; Gaps 1;  
 QY 1 CACAAAGTCACACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60  
 Db 4888 CACAAAGTCACACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 4947  
 QY 61 ACCCAACATACCTCACTCAGAGTCTCAGTTGGAATCAGTTCTTCCAAATCTGTGAA 120  
 Db 4948 ATCACCACCTTACTTACTCAAGTCCCAAGAGGGTTAGTCTATCCAAACCACTACA 5007  
 QY 121 ATTACATCAACACCAACCACTACTAGTTTCACACACACAGAGTCAAGTCAAACTTG 180  
 Db 5008 ACCATCATCAATCCACCAATTCAGCCACACATCATCCCAACCAAAAGTCAGAAACA 5067  
 QY 181 CAAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240  
 Db 5068 CACCAACACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 5127  
 QY 241 ACTACAAAACACGCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 16:51:32 ; Search time 2923.85 Seconds  
(without alignments)  
10004.074 Million cell updates/sec

Title: US-09-462-816-3  
Perfect score: 715  
Sequence: 1 cacaaagtcacactaacac.....gtagttattataaaaaaaa 715

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_ov.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_man.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID       | Description        |
|------------|-------|---------------|--------|----|----------|--------------------|
| 1          | 715   | 100.0         | 715    | 6  | BD081934 | BD081934 Nucleic a |
| 2          | 715   | 100.0         | 920    | 6  | A16257   | A16257 Synthetic R |
| 3          | 715   | 100.0         | 920    | 6  | A16258   | A16258 Synthetic R |
| 4          | 715   | 100.0         | 920    | 6  | AR080406 | AR080406 Sequence  |
| 5          | 715   | 100.0         | 920    | 6  | AR092530 | AR092530 Sequence  |
| 6          | 715   | 100.0         | 920    | 6  | AR122885 | AR122885 Sequence  |
| 7          | 715   | 100.0         | 920    | 6  | AR123540 | AR123540 Sequence  |
| 8          | 715   | 100.0         | 920    | 6  | AR148357 | AR148357 Sequence  |
| 9          | 715   | 100.0         | 920    | 6  | BD081933 | BD081933 Nucleic a |
| 10         | 710   | 99.3          | 923    | 14 | HRSRNAG  | X17085 Human respi |
| 11         | 709   | 99.2          | 922    | 14 | HRSVGL16 | Z33429 Human respi |
| 12         | 699.2 | 97.8          | 917    | 14 | RSHGLYG  | M17212 Human respi |
| 13         | 696   | 97.3          | 894    | 6  | AR080424 | AR080424 Sequence  |
| 14         | 696   | 97.3          | 894    | 6  | AR092548 | AR092548 Sequence  |
| 15         | 654.8 | 91.6          | 8510   | 14 | RSHICE   | M11486 Human respi |
| 16         | 654.8 | 91.6          | 15222  | 6  | AR093219 | AR093219 Sequence  |
| 17         | 654.8 | 91.6          | 15222  | 6  | BD091237 | BD091237 RNase L a |
| 18         | 654.8 | 91.6          | 15222  | 14 | HRU50362 | U50362 Human respi |
| 19         | 654.8 | 91.6          | 15222  | 14 | HRU50363 | U50363 Human respi |
| 20         | 654.8 | 91.6          | 15222  | 14 | HRU63644 | U63644 Human respi |
| 21         | 654.8 | 91.6          | 15222  | 14 | RSHSEQ   | M74568 Human respi |
| 22         | 654.8 | 91.6          | 15223  | 6  | AR089137 | AR089137 Sequence  |
| 23         | 654.8 | 91.6          | 15223  | 14 | AF035006 | AF035006 Human res |
| 24         | 649.6 | 90.9          | 914    | 14 | PARSENVG | X03149 Respiratory |
| 25         | 649.6 | 90.9          | 918    | 6  | AX047075 | AX047075 Sequence  |
| 26         | 646.2 | 90.4          | 897    | 6  | AX339011 | AX339011 Sequence  |
| 27         | 641.8 | 89.8          | 15190  | 14 | RSU39661 | U39662 Human respi |
| 28         | 641.8 | 89.8          | 15191  | 14 | RSU39661 | U39661 Respiratory |
| 29         | 628.6 | 87.9          | 840    | 14 | AF193308 | AF193308 Human res |
| 30         | 628.2 | 87.9          | 918    | 14 | AF065406 | AF065406 Human res |
| 31         | 625.4 | 87.5          | 840    | 14 | AF193306 | AF193306 Human res |
| 32         | 621   | 86.9          | 922    | 14 | HRSVGL12 | Z33412 Human respi |
| 33         | 621   | 86.9          | 922    | 14 | HRSVGL11 | Z33456 Human respi |
| 34         | 621   | 86.9          | 922    | 14 | HRSVGL11 | Z33412 Human respi |
| 35         | 621   | 86.9          | 922    | 14 | HRSVGL11 | Z33456 Human respi |
| 36         | 620.8 | 86.8          | 887    | 14 | AF065409 | AF065409 Human res |
| 37         | 620.2 | 86.7          | 918    | 14 | AF065405 | AF065405 Human res |
| 38         | 619.4 | 86.6          | 922    | 14 | HRSVGL2  | Z33415 Human respi |
| 39         | 618.4 | 86.5          | 922    | 14 | HRSVGL1  | Z33410 Human respi |
| 40         | 618.4 | 86.5          | 922    | 14 | HRSVGL1  | Z33454 Human respi |
| 41         | 618.4 | 86.5          | 922    | 14 | HRSVGL1  | Z33454 Human respi |
| 42         | 617.8 | 86.4          | 922    | 14 | AF065256 | AF065256 Human res |
| 43         | 617.8 | 86.4          | 922    | 14 | AF065258 | AF065258 Human res |
| 44         | 617.8 | 86.4          | 922    | 14 | HRSVGL13 | Z33426 Human respi |
| 45         | 617.8 | 86.4          | 922    | 14 | HRSVGL9  | Z33422 Human respi |

ALIGNMENTS

RESULT 1  
BD081934  
LOCUS BD081934 715 bp DNA linear PAT 27-AUG-2002  
DEFINITION Nucleic acid vaccines encoding G protein of respiratory syncytial virus.  
ACCESSION BD081934  
VERSION BD081934.1 GI:22627544  
KEYWORDS JP 2001512662-A/2  
SOURCE Respiratory syncytial virus  
ORGANISM Respiratory syncytial virus  
Viruses; ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE 1 (bases 1 to 715)  
AUTHORS Li,X.; Sambhara,S. and Klein,M.H.  
TITLE Nucleic acid vaccines encoding G protein of respiratory syncytial

JOURNAL Patent: JP 2001512662-A 2 28-AUG-2001;  
CONNAUGHT LABORATORIES LTD  
COMMENT Respiratory syncytial virus  
PD JP 2001512662-A/2  
PN 28-AUG-2001  
PF 16-JUL-1998 JP 2000503216  
PR 18-JUL-1997 US 08/896442  
FI XIOMAO LI,SURYAPRAKESH SAMBHARA,MICHEL H KLEIN PC  
C12N15/09,A61K48/00,C07K16/10,C12P21/08,C12Q1/68,G01N33/53, PC  
C12N15/00  
CC Nucleic acid vaccines encoding G protein of respiratory CC  
syncytial virus  
FH Key Location/Qualifiers  
FT CDS (1)..(699).  
FEATURES  
source  
1..715  
/organism="Respiratory syncytial virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:12814"  
BASE COUNT 305 a 240 c 70 g 100 t  
ORIGIN  
Query Match 100.0%; Score 715; DB 6; Length 715;  
Best Local Similarity 100.0%; Pred. No. 3.3e-151;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CACAAAGTCACACTAAGTCACTGCAATCATACAGATGCAACGCGAGATCAAGAACACA 60  
Db 1 CACAAAGTCACACTAAGTCACTGCAATCATACAGATGCAACGCGAGATCAAGAACACA 60  
Qy 61 ACCCCAAATACCTCAGTCACTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 120  
Db 61 ACCCCAAATACCTCAGTCACTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 120  
Qy 121 ATTACATCAACAAACCCACCATCTAGCTTCAACCAACACGAGGAGTCAAGTCAAACTG 180  
Db 121 ATTACATCAACAAACCCACCATCTAGCTTCAACCAACACGAGGAGTCAAGTCAAACTG 180  
Qy 181 CAACCCCAAGTCAGTCAAGTCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
Db 181 CAACCCCAAGTCAGTCAAGTCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
Qy 241 ACTACAAAACACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 241 ACTACAAAACACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Qy 301 GTGTTTAACTTTGTACCTGAGCATATGCAAGCAACCAATCCAACTGCTGGGCTATCTGC 360  
Db 301 GTGTTTAACTTTGTACCTGAGCATATGCAAGCAACCAATCCAACTGCTGGGCTATCTGC 360  
Qy 361 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCA 420  
Db 361 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCA 420  
Qy 421 ACCTTCAAGCAACCAAAAGATCTCAACCTCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
Db 421 ACCTTCAAGCAACCAAAAGATCTCAACCTCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
Qy 481 ACCACCAAGCCCAAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
Db 481 ACCACCAAGCCCAAGAGGCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
Qy 541 CTGCTCACCACCAACCAAGGAAATCCAAAATCTCAAGTCAAAATGGAACCTTCCAC 600  
Db 541 CTGCTCACCACCAACCAAGGAAATCCAAAATCTCAAGTCAAAATGGAACCTTCCAC 600  
Qy 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAAATCGGAGCACCA 660  
Db 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAAATCGGAGCACCA 660  
Qy 661 TCACAAACCTCATCTCCACCCCAACCAACGAGGAGTCTTATTAACCAACCAACCAACCAACCA 715  
Db 661 TCACAAACCTCATCTCCACCCCAACCAACGAGGAGTCTTATTAACCAACCAACCAACCAACCA 715

RESULT 2  
A16257  
LOCUS Synthetic RSV G gene (seq ID No: 7). 920 bp DNA linear PAT 03-OCT-1994  
DEFINITION  
A16257  
ACCESSION  
A16257.1 GI:640933  
KEYWORDS  
synthetic construct  
synthetic construct  
artificial sequences.  
SOURCE  
1 (bases 1 to 920)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1..920  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
8..904  
/gene="RSV G"  
8..904  
/gene="RSV G"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA01273.1"  
/db\_xref="GI:640934"  
/translation="MSKNKDQRTAKLEKTLNHLFISSGLYKLNLSVAQITLS  
ILAMISLTITAIPIFIASANKVLTITAIQDTSQIKNTPTVLTITDQPGIGIFS  
NUSEITSQITTLASTTPGKSNLOPVTNTTTPQSPKPTTKQKQNPKNPN  
NDFHVFVPCISCSNRPCTWAICRIPIKPKPKKTTTKPKTFKTKDLKPK  
TTPKEVPPTKPTPEPTINTKINITITLLTNTTGNPKLTSQMETFHSTSSBGNLSP  
SQVSTSEHPSQSPSPNTTQ"  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN  
Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3.3e-151;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CACAAAGTCACACTAAGTCACTGCAATCATACAGATGCAACGCGAGATCAAGAACACA 60  
Db 206 CACAAAGTCACACTAAGTCACTGCAATCATACAGATGCAACGCGAGATCAAGAACACA 265  
Qy 61 ACCCCAAATACCTCAGTCACTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 120  
Db 266 ACCCCAAATACCTCAGTCACTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 325  
Qy 121 ATTACATCAAAACCCACCATCTAGCTTCAACCAACCAACCAACCAACCAACCAACCAACCTG 180  
Db 326 ATTACATCAAAACCCACCATCTAGCTTCAACCAACCAACCAACCAACCAACCAACCAACCTG 385  
Qy 181 CAACCCCAAGTCAGTCAAGCTAATAAACAACAACAACCAACCAACCAACCAACCAACCAACCA 240  
Db 386 CAACCCCAAGTCAGTCAAGCTAATAAACAACAACAACCAACCAACCAACCAACCAACCAACCA 445  
Qy 241 ACTACAAAACACGCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 446 ACTACAAAACACGCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
Qy 301 GTGTTTAACTTTGTACCTGAGCATATGCAAGCAACCAATCCAACTGCTGGGCTATCTGC 360  
Db 506 GTGTTTAACTTTGTACCTGAGCATATGCAAGCAACCAATCCAACTGCTGGGCTATCTGC 565  
Qy 361 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCA 420  
Db 566 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCA 625  
Qy 421 ACCTTCAAGCAACCAAAAGATCTCAACCTCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
Db 626 ACCTTCAAGCAACCAAAAGATCTCAACCTCAACCAACCAACCAACCAACCAACCAACCAACCA 685



QY 481 ACCACCAAGCCACAGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 540  
Db 686 ACCACCAAGCCACAGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 745  
QY 541 CTGCTCAGCAACACACAGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 600  
Db 746 CTGCTCAGCAACACACAGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 805  
QY 601 TCACACCTCTCCGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 660  
Db 806 TCACACCTCTCCGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 865  
QY 661 TCACACCTCTCTCCGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 715  
Db 866 TCACACCTCTCTCCGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 920

RESULT 3  
Al6258/c  
LOCUS Synthetic RSV G gene (seq ID No: 7) linear PAT 03-OCT-1994  
ACCESSION Al6258  
VERSION Al6258.1 GI:640935  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS CHIMERIC IMMUNOGENS  
TITLE Patent: WO 93/4207-A 27 22-JUL-1993;  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..920  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

BASE COUNT 155 a 95 c 290 g 380 t  
ORIGIN  
Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3.3e-151;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAAGTCGCAATCATCAAGATGCAACAGCCGAGATCAAGAACACA 60  
Db 715 CACAAAGTCACACTAAGTCGCAATCATCAAGATGCAACAGCCGAGATCAAGAACACA 656  
QY 61 ACCCCCAACATACCTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 120  
Db 655 ACCCCCAACATACCTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 596  
QY 121 ATTACATCAAAACCAACCATCTAGCTTCAACACACAGGAGTCAAGTCAAAACCTG 180  
Db 595 ATTACATCAAAACCAACCATCTAGCTTCAACACACAGGAGTCAAGTCAAAACCTG 536  
QY 181 CAACCCCAACAGTCAGTCAACACACACACACACACACACACACACACACACACACAC 240  
Db 535 CAACCCCAACAGTCAGTCAACACACACACACACACACACACACACACACACACACAC 476  
QY 241 ACTACAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 475 ACTACAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 416  
QY 301 GTGTTTAACTTTGTACCTTGAGCATATGAGCAACCAATCAACCTGCTGGGCTATCTGC 360  
Db 415 GTGTTTAACTTTGTACCTTGAGCATATGAGCAACCAATCAACCTGCTGGGCTATCTGC 356  
QY 361 AAAAGAAATACCAAAACCAAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
Db 355 AAAAGAAATACCAAAACCAAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 296  
QY 421 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480

Db 295 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 236  
QY 481 ACCACCAAGCCACAGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 540  
Db 235 ACCACCAAGCCACAGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 176  
QY 541 CTGCTCAGCAACACACAGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 600  
Db 175 CTGCTCAGCAACACACAGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 116  
QY 601 TCACACCTCTCCGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 660  
Db 115 TCACACCTCTCCGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 56  
QY 661 TCACACCTCTCTCCGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 715  
Db 55 TCACACCTCTCTCCGAGAGCCAAACCATCAACACCCAAACCAACATCAACACTTACA 1

RESULT 4  
AR080406  
LOCUS Sequence 7 from patent US 5968776 linear PAT 31-AUG-2000  
DEFINITION AR080406  
ACCESSION AR080406  
VERSION AR080406.1 GI:10007141  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewaszyshyn, M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
JOURNAL Patent: US 5968776-A 7 19-OCT-1999;  
FEATURES  
source 1..920  
/organism="unknown"

BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN  
Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3.3e-151;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAAGTCGCAATCATCAAGATGCAACAGCCGAGATCAAGAACACA 60  
Db 206 CACAAAGTCACACTAAGTCGCAATCATCAAGATGCAACAGCCGAGATCAAGAACACA 265  
QY 61 ACCCCCAACATACCTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 120  
Db 266 ACCCCCAACATACCTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 325  
QY 121 ATTACATCAAAACCAACCATCTAGCTTCAACACACAGGAGTCAAGTCAAAACCTG 180  
Db 326 ATTACATCAAAACCAACCATCTAGCTTCAACACACAGGAGTCAAGTCAAAACCTG 385  
QY 181 CAACCCCAACAGTCAGTCAACACACACACACACACACACACACACACACACACACAC 240  
Db 386 CAACCCCAACAGTCAGTCAACACACACACACACACACACACACACACACACACACAC 445  
QY 241 ACTACAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 446 ACTACAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
QY 301 GTGTTTAACTTTGTACCTTGAGCATATGAGCAACCAATCAACCTGCTGGGCTATCTGC 360  
Db 506 GTGTTTAACTTTGTACCTTGAGCATATGAGCAACCAATCAACCTGCTGGGCTATCTGC 565  
QY 361 AAAAGAAATACCAAAACCAAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
Db 566 AAAAGAAATACCAAAACCAAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 625



Db 566 AAAAGATACCAACCAAAACAGGAAAGAAACCAACCAAGCTTACAAAAACCA 625  
Qy 421 ACCTTCAAGACAAACCAAAAGATCTCAAACTCAAACTTAAACCAAGGAGTACCC 480  
Db 626 ACCTTCAAGACAAACCAAAAGATCTCAAACTTAAACCAAGGAGTACCC 685  
Qy 481 ACCACCAAGCCACAGAGAGCCAAACCATCAACACCAACCAAAACATCACTACTACA 540  
Db 686 ACCACCAAGCCACAGAGAGCCAAACCATCAACACCAACCAAAACATCACTACTACA 745  
Qy 541 CTGCTCACAACACACACAGAGAAATCCAAAATCAAACTCAAACTCAAACTCAAACTCA 600  
Db 746 CTGCTCACAACACACAGAGAAATCCAAAATCAAACTCAAACTCAAACTCAAACTCA 805  
Qy 601 TCAACCTCTCCGAGGAAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGCACC 660  
Db 806 TCAACCTCTCCGAGGAAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGCACC 865  
Qy 661 TCAACCTCTCTCCACCCCAACACACAGCCAGTCTTATTAATAAAAAA 715  
Db 866 TCAACCTCTCTCCACCCCAACACACAGCCAGTCTTATTAATAAAAAA 920

RESULT 7  
ARI123540  
LOCUS ARI123540 920 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 7 from patent US 6171783.  
ACCESSION ARI123540  
VERSION ARI123540.1 GI:14108901  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewasyszyn, M.E.  
TITLE Infection detection method using chimeric protein  
JOURNAL Patent: US 6171783-A 7 09-JAN-2001;  
FEATURES Location/Qualifiers  
source 1..920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN  
Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3.3e-151;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CACAAGTCACTTAACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60  
Db 206 CACAAGTCACTTAACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 265  
Qy 61 ACCCCACATACCTCACTCAGGATCTCAGTTGGATCAGTTCTCAATCTGTCTGAA 120  
Db 266 ACCCCACATACCTCACTCAGGATCTCAGTTGGATCAGTTCTCAATCTGTCTGAA 325  
Qy 121 ATTACATCAAAACCAACCACTACTAGTCTCAACACACAGAGTCAAGTCAAACTG 180  
Db 326 ATTACATCAAAACCAACCACTACTAGTCTCAACACACAGAGTCAAGTCAAACTG 385  
Qy 181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240  
Db 386 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 445  
Qy 241 ACTACAAACCAAGCCCAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 446 ACTACAAACCAAGCCCAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
Qy 301 GTGTTTAACTTTGTACCTGCAGATATGAGCAACCAACCAACCAACCAACCAACCA 360  
Db 506 GTGTTTAACTTTGTACCTGCAGATATGAGCAACCAACCAACCAACCAACCAACCA 565  
Qy 361 AAAAGATACCAACCAAAACAGGAAAGAAACCAACCAAGCTTACAAAAACCA 420

Db 566 AAAAGATACCAACCAAAACAGGAAAGAAACCAACCAAGCTTACAAAAACCA 625  
Qy 421 ACCTTCAAGACAAACCAAAAGATCTCAAACTCAAACTTAAACCAAGGAGTACCC 480  
Db 626 ACCTTCAAGACAAACCAAAAGATCTCAAACTTAAACCAAGGAGTACCC 685  
Qy 481 ACCACCAAGCCACAGAGAGCCAAACCATCAACACCAACCAAAACATCACTACTACA 540  
Db 686 ACCACCAAGCCACAGAGAGCCAAACCATCAACACCAACCAAAACATCACTACTACA 745  
Qy 541 CTGCTCACAACACACAGAGAAATCCAAAATCAAACTCAAACTCAAACTCAAACTCA 600  
Db 746 CTGCTCACAACACACAGAGAAATCCAAAATCAAACTCAAACTCAAACTCAAACTCA 805  
Qy 601 TCAACCTCTCCGAGGAAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGCACC 660  
Db 806 TCAACCTCTCCGAGGAAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGCACC 865  
Qy 661 TCAACCTCTCTCCACCCCAACACACAGCCAGTCTTATTAATAAAAAA 715  
Db 866 TCAACCTCTCTCCACCCCAACACACAGCCAGTCTTATTAATAAAAAA 920

RESULT 8  
ARI148357  
LOCUS ARI148357 920 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 7 from patent US 6225091.  
ACCESSION ARI148357  
VERSION ARI148357.1 GI:15112447  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewasyszyn, M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
JOURNAL Patent: US 6225091-A 7 01-MAY-2001;  
FEATURES Location/Qualifiers  
source 1..920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN  
Query Match 100.0%; Score 715; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3.3e-151;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CACAAGTCACTTAACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60  
Db 206 CACAAGTCACTTAACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 265  
Qy 61 ACCCCACATACCTCACTCAGGATCTCAGTTGGATCAGTTCTCAATCTGTCTGAA 120  
Db 266 ACCCCACATACCTCACTCAGGATCTCAGTTGGATCAGTTCTCAATCTGTCTGAA 325  
Qy 121 ATTACATCAAAACCAACCACTACTAGTCTCAACACACAGAGTCAAGTCAAACTG 180  
Db 326 ATTACATCAAAACCAACCACTACTAGTCTCAACACACAGAGTCAAGTCAAACTG 385  
Qy 181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240  
Db 386 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 445  
Qy 241 ACTACAAACCAAGCCCAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 446 ACTACAAACCAAGCCCAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
Qy 301 GTGTTTAACTTTGTACCTGCAGATATGAGCAACCAACCAACCAACCAACCAACCA 360  
Db 506 GTGTTTAACTTTGTACCTGCAGATATGAGCAACCAACCAACCAACCAACCAACCA 565

|                       |     |   |     |
|-----------------------|-----|---|-----|
| QY                    | 361 | AAAAGAATACCAACCAAAAAACCAGGAAAGAAAAACCCACCAAGCCTTACAAAAAACCA   | 420 |
| Db                    | 566 | AAAAGAATACCAACCAAAAAACCAGGAAAGAAAAACCCACCAAGCCTTACAAAAAACCA   | 625 |
| QY                    | 421 | ACCTTCAAGACACCAAAAAAAGATCTCAAACCTCAAGCTCAACCACTTAACCAAGGAAGTACCC  | 480 |
| Db                    | 626 | ACCTTCAAGACACCAAAAAAAGATCTCAAACCTCAAGCTCAACCACTTAACCAAGGAAGTACCC  | 685 |
| QY                    | 481 | ACCACCAAGGCCACAGAAAGCCCAACATCAACCAACCAACCAACCAACCAACTACA  | 540 |
| Db                    | 686 | ACCACCAAGGCCACAGAAAGCCCAACATCAACCAACCAACCAACCAACTACA  | 745 |
| QY                    | 541 | CTGCCTCACCAACCAACCAACCAAGGAATTCAAAACCTCAAACTCAAAATGGAAACCTTCAC  | 600 |
| Db                    | 746 | CTGCCTCACCAACCAACCAACCAAGGAATTCAAAACCTCAAACTCAAAATGGAAACCTTCAC  | 805 |
| QY                    | 601 | TCAACCTCTCCGNAAGCGAATCTTAAGCCCTTCTCAAGTCTCAACCACTCCGAGACCCA   | 660 |
| Db                    | 806 | TCAACCTCTCCGNAAGCGAATCTTAAGCCCTTCTCAAGTCTCAACCACTCCGAGACCCA   | 865 |
| QY                    | 661 | TCAACACCCCTCATCTCCACCCCAACCAACCAAGCCAGCTAGTTATTATAAAAAAAAAA   | 715 |
| Db                    | 866 | TCAACACCCCTCATCTCCACCCCAACCAACCAAGCCAGCTAGTTATTATAAAAAAAAAA   | 920 |
| RESULT 9              |     |   |     |
| BD081933              |     |   |     |
| LOCUS                 |     | 920 bp DNA linear PAT 27-AUG-2002   |     |
| DEFINITION            |     | Nucleic acid vaccines encoding G protein of respiratory syncytial virus.  |     |
| ACCESSION             |     | BD081933  |     |
| VERSION               |     | BD081933.1 GI:22627543  |     |
| KEYWORDS              |     | JP 2001512662-A/1.  |     |
| SOURCE                |     | Respiratory syncytial virus   |     |
| ORGANISM              |     | Respiratory syncytial virus   |     |
| REFERENCE             |     | Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.   |     |
| AUTHORS               |     | Li, X., Sambhara, S. and Klein, M.H.  |     |
| TITLE                 |     | Nucleic acid vaccines encoding G protein of respiratory syncytial   |     |
| JOURNAL               |     | Patent: JP 2001512662-A 1 28-AUG-2001;  |     |
| COMMENT               |     | CONNAUGHT LABORATORIES LTD<br>OS Respiratory syncytial virus<br>PN JP 2001512662-A/1<br>PD 28-AUG-2001<br>PF 16-JUL-1998 JP 2000503216<br>PR 18-JUL-1997 US 087896442<br>PI XIAOMAO LI,SURYAPRAKESH SAMBHARA,MICHEL H KLEIN PC<br>C12N15/09,A61K48/00,C07K16/10,C12P21/08,C12Q1/68,G01N33/53, PC<br>C12N15/09 |     |
| FEATURES              |     | CC Nucleic acid vaccines encoding G protein of respiratory CC<br>syncytial virus  |     |
| source                |     | FH Key Location/Qualifiers<br>FT CDS (7)..(901).  |     |
|                       |     | 1..920<br>/organism="Respiratory syncytial virus"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:12814"  |     |
| BASE COUNT            |     | 380 a 290 c 95 g 155 t  |     |
| ORIGIN                |     |   |     |
| Query Match           |     | 100.0%; Score 715; DB 6; Length 920;  |     |
| Best Local Similarity |     | 100.0%; Pred. No. 3.3e-151;   |     |
| Matches               |     | 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  |     |
| QY                    | 1   | CACAAAGTCACCTAAACCTGCAATCATACAGATGCAACAGCAGATCAAGACACA  | 60  |
| Db                    | 206 | CACAAAGTCACCTAAACCTGCAATCATACAGATGCAACAGCAGATCAAGACACA  | 265 |
| QY                    | 61  | ACCCCAACATACCTCATTCAAGATCCTCAGCTTGGAATCAGCTTCTCCAAATCTGTCTGAA   | 120 |

|   |  |  |  |
|---|--|--|--|
| /db_xref="SPTREMBL:Q01929"  |  | Viruses; serNA negative-strand viruses; Mononegavirales;                 |  |
| /translation="MSKNDQRTAKTLEKWTDLNLLFISSGLYKLNLSIAQITLS                      |  | Paramyxoviridae; Pneumovirinae; Pneumovirus.                             |  |
| ILAMIISTLIITAIIFIASANKHVTITAIIOQATSOIKNTPTLYTQDPLGISFS                      |  | 1 (bases 1 to 922)   |  |
| NLSEITSTVITLSTTPGVSNLQPTTKNTTKNTTQPSKPTTKQRQKNPKPN                          |  | Garcia, O., Martin, M., Dopazo, J., Arbiza, J., Fabraile, S., Russi, J., |  |
| NDPHEVFNPCISCSNNPCWAICRIIPNKKPGKTTTKPKPTFKTKDLKPKQ                          |  | Hortal, M., Perez-Brena, P., Martinez, I., Garcia-Barreno, B. and        |  |
| TKPEVPTTKPTTEPTINTTKNTITLITLNTNNTGNPKLTSQMETFHTSSGNLSP                      |  | Melero, J.A.   |  |
| SQVSTTSEHPSPSPSPNTTQ"   |  | Evolutionary pattern of human respiratory syncytial virus (subgroup      |  |
| BASE COUNT 379 a 290 c 98 g 156 t   |  | A): cocirculating lineages and correlation of genetic and antigenic      |  |
| ORIGIN  |  | Changes in the G glycoprotein  |  |
| Query Match 99.3%; Score 710; DB 14; Length 923;                            |  | J. Virol. 68 (9), 5448-5459 (1994)                                       |  |
| Best Local Similarity 100.0%; Pred. No. 4.4e-150;                           |  | 8057427  |  |
| Matches 710; Conservative 0; Mismatches 0; Indels 0; Gaps 0;                |  | 2 (bases 1 to 922)   |  |
| Qy 1 CACAAAGTCACACTAACCACTGCATCATACAGATGCAACAGCCAGATCAAGAACACA 60           |  | Dopazo, J.   |  |
| Db 214 CACAAAGTCACACTAACCACTGCATCATACAGATGCAACAGCCAGATCAAGAACACA 273        |  | Direct Submission  |  |
| Qy 61 ACCCAACATACCTCAGTACAGATCTCAGTTTGGATTCAGTTTCCAAATCTGTCTGAA 120         |  | Submitted (10-MAY-1994) Dopazo J., Centro Nacional de Biotecnologia      |  |
| Db 274 ACCCAACATACCTCAGTACAGATCTCAGTTTGGATTCAGTTTCCAAATCTGTCTGAA 333        |  | - CSIC, Bioinformatica, Universidad Autonoma, Cantoblanco, Madrid,       |  |
| Qy 121 ATTACATCAAAACCAACCACTACCTAGCTTCAACCAACCAAGAGTCAAGTCAAACTGT 180       |  | SPAIN, 28049   |  |
| Db 334 ATTACATCAAAACCAACCACTACCTAGCTTCAACCAACCAAGAGTCAAGTCAAACTGT 393       |  | Location/Qualifiers  |  |
| Qy 181 CAACCCACACAGTCAAGCTAAACACACACACACACACACACACACACACACACACAC 240        |  | 1..922   |  |
| Db 394 CAACCCACACAGTCAAGCTAAACACACACACACACACACACACACACACACACACAC 453        |  | /organism="Human respiratory syncytial virus"                            |  |
| Qy 241 ACTACAAAAC 300         |  | /mol_type="genomic RNA"  |  |
| Db 454 ACTACAAAAC 513         |  | /strain="subgroup A"   |  |
| Qy 301 GTGTTTAACTTTGTACCTTGCAGCATATGCAGCAACAATCCAACTCTGGGCTATCTGC 360       |  | /isolate="MON 7 91 (Montevideo/Uruguay, 1991)"                           |  |
| Db 514 GTGTTTAACTTTGTACCTTGCAGCATATGCAGCAACAATCCAACTCTGGGCTATCTGC 573       |  | /db_xref="taxon:11250"   |  |
|   |  | /cell_line="HEp-2"   |  |
|   |  | 16..912  |  |
|   |  | /gene="G"  |  |
|   |  | 16..912  |  |
|   |  | /partial   |  |
|   |  | /genes="G"   |  |
|   |  | /codon_start=1   |  |
|   |  | /product="glycoprotein"  |  |
|   |  | /protein_id="CAA83872.1"   |  |
|   |  | /db_xref="GI:485889"   |  |
|   |  | /db_xref="SPTREMBL:Q82068"   |  |
|   |  | /translations="ISKNDQRTAKTLEKWTDLNLLFISSGLYKLNLSIAQITLS                  |  |
|   |  | ILAMIISTLIITAIIFIASANKHVTITAIIOQATSOIKNTPTLYTQDPLGISFS                   |  |
|   |  | NLSEITSTVITLSTTPGVSNLQPTTKNTTKNTTQPSKPTTKQRQKNPKPN                       |  |
|   |  | NDPHEVFNPCISCSNNPCWAICRIIPNKKPGKTTTKPKPTFKTKDLKPKQ                       |  |
|   |  | TKPEVPTTKPTTEPTINTTKNTITLITLNTNNTGNPKLTSQMETFHTSSGNLSP                   |  |
|   |  | SQVSTTSEHPSPSPSPNTTQ"  |  |
| BASE COUNT 378 a 291 c 97 g 156 t   |  |  |  |
| ORIGIN  |  |  |  |
| Query Match 99.2%; Score 709; DB 14; Length 922;                            |  |  |  |
| Best Local Similarity 100.0%; Pred. No. 7.4e-150;                           |  |  |  |
| Matches 709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;                |  |  |  |
| Qy 1 CACAAAGTCACACTAACCACTGCATCATACAGATGCAACAGCCAGATCAAGAACACA 60           |  |  |  |
| Db 214 CACAAAGTCACACTAACCACTGCATCATACAGATGCAACAGCCAGATCAAGAACACA 273        |  |  |  |
| Qy 61 ACCCAACATACCTCAGTACAGATCTCAGTTTGGATTCAGTTTCCAAATCTGTCTGAA 120         |  |  |  |
| Db 274 ACCCAACATACCTCAGTACAGATCTCAGTTTGGATTCAGTTTCCAAATCTGTCTGAA 333        |  |  |  |
| Qy 121 ATTACATCAAAACCAACCACTACCTAGCTTCAACCAACCAAGAGTCAAGTCAAACTGT 180       |  |  |  |
| Db 334 ATTACATCAAAACCAACCACTACCTAGCTTCAACCAACCAAGAGTCAAGTCAAACTGT 393       |  |  |  |
| Qy 181 CAACCCACACAGTCAAGCTAAACACACACACACACACACACACACACACACACACAC 240        |  |  |  |
| Db 394 CAACCCACACAGTCAAGCTAAACACACACACACACACACACACACACACACACACAC 453        |  |  |  |
| Qy 241 ACTACAAAAC 300         |  |  |  |
| Db 454 ACTACAAAAC 513         |  |  |  |
| Qy 301 GTGTTTAACTTTGTACCTTGCAGCATATGCAGCAACAATCCAACTCTGGGCTATCTGC 360       |  |  |  |
| Db 514 GTGTTTAACTTTGTACCTTGCAGCATATGCAGCAACAATCCAACTCTGGGCTATCTGC 573       |  |  |  |
| Qy 361 AAAAAAGTACCAACAAAAACACAGGAAGAAAAACCAACCAACCAACCAACCAACCAACCA 420     |  |  |  |
| Db 574 AAAAAAGTACCAACAAAAACACAGGAAGAAAAACCAACCAACCAACCAACCAACCAACCA 633     |  |  |  |
| Qy 421 ACCTTCAAGCAACCAACAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAC 480    |  |  |  |
| Db 634 ACCTTCAAGCAACCAACAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAC 693    |  |  |  |
| Qy 481 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCAAC 540     |  |  |  |
| Db 694 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCAAC 753     |  |  |  |
| Qy 541 CTGCTCAACCAACCAACCAACCAAGAAATCCAAACTCAAACTCAAACTCAAACTCAAACTCCAC 600 |  |  |  |
| Db 754 CTGCTCAACCAACCAACCAACCAAGAAATCCAAACTCAAACTCAAACTCAAACTCAAACTCCAC 813 |  |  |  |
| Qy 601 TCAACCTCTCCAGAGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCAACTCCAGGACCCCA 660    |  |  |  |
| Db 814 TCAACCTCTCCAGAGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCAGGACCCCA 873          |  |  |  |
| Qy 661 TCACACCTCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 710    |  |  |  |
| Db 874 TCACACCTCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 923    |  |  |  |
| RESULT 11   |  |  |  |
| HRVGL16   |  |  |  |
| LOCUS   |  |  |  |
| DEFINITION  |  | HRVGL16 922 bp RNA linear VRL 05-JUN-1997                                |  |
|   |  | Human respiratory syncytial virus (MON-7-91) subgroup A, G gene for      |  |
|   |  | glycoprotein.  |  |
| ACCESSION   |  | 233429   |  |
| VERSION   |  | 233429.1 GI:485888   |  |
| KEYWORDS  |  | G gene; glycoprotein.  |  |
| SOURCE  |  | Human respiratory syncytial virus  |  |
| ORGANISM  |  | Human respiratory syncytial virus  |  |

QY 361 AAAAGAATACCAACAAAAACCGAGAAAGAAAACCAACCAAGCCTTACAAAAAACCA 420  
 DB 574 AAAAGAATACCAACAAAAACCGAGAAAGAAAACCAACCAAGCCTTACAAAAAACCA 633  
 QY 421 ACCTTCAAGACACCAAAAAAGATCTCAACCTCAACCACTAAACCAAGAAAGTACCC 480  
 DB 634 ACCTTCAAGACACCAAAAAAGATCTCAACCTCAACCACTAAACCAAGAAAGTACCC 693  
 QY 481 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACTACA 540  
 DB 694 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACTACA 753  
 QY 541 CTGCTCAACCAACCAACCAAGCAATCAAAATCTCAACAGTCAAAATGGAACCTTCCAC 600  
 DB 754 CTGCTCAACCAACCAACCAAGCAATCAAAATCTCAACAGTCAAAATGGAACCTTCCAC 813  
 QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAATCCGAGACCCA 660  
 DB 814 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAATCCGAGACCCA 873  
 QY 661 TCACAACTCTATCTCCACCAACCAACACACACGCGAGTAGTATTAAAAA 709  
 DB 874 TCACAACTCTATCTCCACCAACCAACACACGCGAGTAGTATTAAAAA 922

RESULT 12  
 RSHGLYG 917 bp ss-RNA linear VRL 03-AUG-1993  
 LOCUS Human respiratory syncytial virus (subgroup A) attachment protein  
 DEFINITION (G) mRNA, complete cds.  
 ACCESSION M17212.1 GI:333940  
 VERSION 1  
 KEYWORDS attachment glycoprotein; surface glycoprotein.  
 SOURCE Human respiratory syncytial virus  
 ORGANISM Human respiratory syncytial virus  
 REFERENCE 1 (bases 1 to 917)  
 AUTHORS Johnson, P.R., Spriggs, M.K., Olmsted, R.A., and Collins, P.L.  
 TITLE The G glycoprotein of human respiratory syncytial viruses of  
 subgroups A and B: extensive sequence divergence between  
 antigenically related proteins  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (16), 5625-5629 (1987)  
 MEDLINE 87289657  
 PUBMED 2441388  
 COMMENT Original source text: Human respiratory syncytial virus (subgroup A,  
 strain Long), cDNA to viral RNA, clones pU63, pU69, and pA3.  
 The exact 5' end of Long G mRNA was not determined.

FEATURES  
 source  
 1. .917  
 Location/Qualifiers  
 /organism="Human respiratory syncytial virus"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:11250"  
 16. .912  
 /note="attachment glycoprotein (G)"  
 /codon\_start=1  
 /protein\_id="AAA47411.1"  
 /db\_xref="GI:333941"  
 /translation="MSKNKQORTANTLEKTWDTLNHLFTSSGLYKLNLSIAQITLS  
 ILAMITSLIITAIPIASANKHVLITTAIQDTSQIKNTTPTLYLTOPQLGISFS  
 NLSBITSTIITLSTPPGVKSLQPTTKVTKNTTQTPQSKPTTKQKQKPNKPN  
 NDFEVEFNPVCSICSNPTTCAIKRIPNKPKGKTTKPTKPTKPTKTKDHPQ  
 TTKPKVPTTKPTTEPTINTKNTIITLLTNTTGNPKLTQSMETPHSTSSGNLSF  
 SQVSTSEHSPQSPSPNTTQ"  
 BASE COUNT 375 a 290 c 97 g 154 t 1 others  
 ORIGIN  
 Query Match 97.8%; Score 699.2; DB 14; Length 917;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-147;  
 Matches 701; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CACAAAGTCACTAACAACTGCAATCATACAGATCAACAGCCAGATCAAGAACCA 60  
 |||||

DB 214 CACAAAGTCACTAACAACTGCAATCATACAGATCAACAGCCAGATCAAGAACCA 273  
 QY 61 ACCCAACATACCTCACTAGGATCTCAGTCTGGAAATCAGTCTTCCAAATCTGTCTGAA 120  
 DB 274 ACCCAACATACCTCACTAGGATCTCAGTCTGGAAATCAGTCTTCCAAATCTGTCTGAA 333  
 QY 121 ATTACATCAACAAACCAACCATACTAGTCTTCAACACACAGAGTCAAGTCAACCTG 180  
 DB 334 ATTACATCAACAAACCAACCATACTAGTCTTCAACACACAGAGTCAAGTCAACCTG 393  
 QY 181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCC 240  
 DB 394 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCC 453  
 QY 241 ACTACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTT 300  
 DB 454 ACTACAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCTT 513  
 QY 301 GTGTTTAACTTTGTACCCCTGCAGCATATGAGCAACCAACCTTCAACCTTCTGGCTATTCG 360  
 DB 514 GTGTTTAACTTTGTACCCCTGCAGCATATGAGCAACCAACCTTCTGGCTATTCG 573  
 QY 361 AAAAGATATCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 420  
 DB 574 AAAAGATATCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 633  
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 480  
 DB 634 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 693  
 QY 481 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCT 540  
 DB 694 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCT 753  
 QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 600  
 DB 754 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 813  
 QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAATCCGAGACCCA 660  
 DB 814 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAATCCGAGACCCA 873  
 QY 661 TCACAACTCTATCTCCACCAACCAACACACACGCGAGTAGTATT 704  
 DB 874 TCACAACTCTATCTCCACCAACCAACACACGCGAGTAGTATT 917

RESULT 13  
 AR080424  
 LOCUS 894 bp DNA linear PAT 31-AUG-2000  
 DEFINITION Sequence 28 from patent US 5968776.  
 ACCESSION AR080424  
 VERSION AR080424.1 GI:10007159  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 894)  
 AUTHORS Klein, M.H., Du, R.-P. and Ewasysghyn, M.R.  
 TITLE Multimeric hybrid gene encoding a chimeric protein which confers  
 protection against parainfluenza virus and respiratory syncytial  
 virus  
 JOURNAL Patent: US 5968776-A 28 19-OCT-1999;  
 FEATURES Location/Qualifiers  
 source 1. .894  
 /organism="unknown"  
 BASE COUNT 364 a 288 c 93 g 149 t  
 ORIGIN  
 Query Match 97.3%; Score 696; DB 6; Length 894;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-147;  
 Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 CACAAAGTCACACTAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60  
 DB 199 CACAAAGTCACACTAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 258  
 QY 61 ACCCCAAATACCTCACTCAGATCCTCAGCTTGAATCAGCTTCTCCAACTCTCTGAA 120  
 DB 259 ACCCCAAATACCTCACTCAGATCCTCAGCTTGAATCAGCTTCTCCAACTCTCTGAA 318  
 QY 121 ATTACATCACAAACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 180  
 DB 319 ATTACATCACAAACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 378  
 QY 181 CAAACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 240  
 DB 379 CAAACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 438  
 QY 241 ACTACAAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 300  
 DB 439 ACTACAAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 498  
 QY 301 GTGTTTAACTTTGTACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 360  
 DB 499 GTGTTTAACTTTGTACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 558  
 QY 361 AAAGATACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 420  
 DB 559 AAAGATACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 618  
 QY 421 ACCTTCAAGCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 480  
 DB 619 ACCTTCAAGCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 678  
 QY 481 ACACCAAGCCACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 540  
 DB 679 ACACCAAGCCACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 738  
 QY 541 CTGCTCACCACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 600  
 DB 739 CTGCTCACCACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 798  
 QY 601 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGACCCA 660  
 DB 799 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGACCCA 858  
 QY 661 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGACCCA 696  
 DB 859 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGACCCA 894

RESULT 14  
 AR092548  
 LOCUS Sequence 28 from patent US 5998169.  
 ACCESSION AR092548  
 VERSION AR092548.1 GI:10019302  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 894)  
 AUTHORS Klein, M.H., Du, R.-P. and Ewaszyshyn, M.E.  
 TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus

JOURNAL Patent: US 5998169-A 28 07-DEC-1999;  
 FEATURES Location/Qualifiers  
 source 1..894  
 /organism="unknown"  
 BASE COUNT 364 a 288 c 93 g 149 t  
 ORIGIN

Query Match 97.3%; Score 696; DB 6; Length 894;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-147;

Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CACAAAGTCACACTAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60  
 DB 199 CACAAAGTCACACTAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 258  
 QY 61 ACCCCAAATACCTCACTCAGATCCTCAGCTTGAATCAGCTTCTCCAACTCTCTGAA 120  
 DB 259 ACCCCAAATACCTCACTCAGATCCTCAGCTTGAATCAGCTTCTCCAACTCTCTGAA 318  
 QY 121 ATTACATCACAAACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 180  
 DB 319 ATTACATCACAAACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 378  
 QY 181 CAAACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 240  
 DB 379 CAAACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 438  
 QY 241 ACTACAAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 300  
 DB 439 ACTACAAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 498  
 QY 301 GTGTTTAACTTTGTACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 360  
 DB 499 GTGTTTAACTTTGTACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 558  
 QY 361 AAAGATACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 420  
 DB 559 AAAGATACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 618  
 QY 421 ACCTTCAAGCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 480  
 DB 619 ACCTTCAAGCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 678  
 QY 481 ACACCAAGCCACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 540  
 DB 679 ACACCAAGCCACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 738  
 QY 541 CTGCTCACCACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 600  
 DB 739 CTGCTCACCACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 798  
 QY 601 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGACCCA 660  
 DB 799 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGACCCA 858  
 QY 661 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGACCCA 696  
 DB 859 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGACCCA 894

## RESULT 15

## RSHICE

## LOCUS

## DEFINITION

## ACCSSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

RSHICE 8510 bp ss-RNA linear VRL 29-NOV-2000  
 Human respiratory syncytial virus nonstructural protein (1C),  
 nonstructural protein (1B), major nucleocapsid (N), phosphoprotein  
 (P), protein (M), 1A (1A), G (G), protein (P) and  
 envelope-associated protein (22K) gene, complete cds.  
 M11486 K01459 K02719 K03348 K03349 M11217 M11244 M11487 M11505  
 M11514 M11631 M12966  
 M11486.1 GI:333925  
 envelope-associated protein; fusion glycoprotein; major  
 nucleocapsid protein; major surface glycoprotein; matrix protein;  
 nonstructural protein; phosphoprotein.  
 Human respiratory syncytial virus  
 Human respiratory syncytial virus  
 Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 1 (bases 1085 to 2220; 5254 to 5544)  
 Elango, N. and Venkatesan, S.  
 Amino acid sequence of human respiratory syncytial virus  
 nucleocapsid protein  
 Nucleic Acids Res. 11 (17), 5941-5951 (1983)



|           |   |
|-----------|---|
| JOURNAL   | Virology 143 (2), 442-451 (1985)  |
| MEDLINE   | 86045905  |
| PUBMED    | 2998021   |
| REFERENCE | 12 (bases 1081 to 2277)   |
| AUTHORS   | Collins,P.L., Anderson,K., Langer,S.J. and Wertz,G.W.   |
| TITLE     | Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus   |
| JOURNAL   | Virology 146 (1), 69-77 (1985)  |
| MEDLINE   | 85301974  |
| PUBMED    | 3839952   |
| REFERENCE | 13 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578)   |
| AUTHORS   | Collins,P.L., Dickens,L.E., Buckler-White,A., Olmsted,R.A., Spriggs,M.K., Camargo,E. and Coelingh,K.V.  |
| TITLE     | Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order   |
| JOURNAL   | Proc. Natl. Acad. Sci. U.S.A. 83 (13), 4594-4598 (1986)   |
| MEDLINE   | 86259643  |
| PUBMED    | 3460060   |
| COMMENT   | Reprint for [2] and clean copy sequence for [12], [5] kindly provided by P.Collins, 21-FEB-1986.<br>Respiratory syncytial virus is a negative-strand RNA pleomorphic enveloped virus of the genus Pneumovirus. The matrix protein has no homology with the matrix proteins of other negative-stranded RNA viruses, implying that RS virus has undergone extensive evolutionary divergence. Two unidentified reading frames which potentially encode proteins were located: one overlaps the matrix protein and the other the F protein. The biological significance of these two reading frames is not clear. The positive strand is shown.   |
| FEATURES  | Location/Qualifiers<br>1..8510<br>/organism="Human respiratory syncytial virus"<br>/mol_type="genomic RNA"<br>/strain="A2"<br>/db_xref="taxon:11250"<br>/lab_host="HEP-2 cell"<br>1..528<br>/gene="1C"<br>1..528<br>/gene="1C"<br>55..474<br>/gene="1C"<br>/codon_start=1<br>/product="nonstructural protein"<br>/protein_id="AAB59850.1"<br>/db_xref="GI:333926"<br>/translation="MGSNSLSMKIVRLQNLQNFQNDVALLKICTYTDKGLHLTNALAKA<br>VHTTKLNGIVFVHTSSDICPNNNIVKSNFTTWPVLQNGYIWEWMLTCHSCP<br>GLDDNCEIKFKSKLSDSTWYMNQLSLLGFDLNP"<br>171<br>/gene="1C"<br>/note="c in [11]; t in [6]"<br>/replace="t"<br>552..1050<br>/gene="1B"<br>552..1050<br>/gene="1B"<br>584..958<br>/gene="1B"<br>/codon_start=1<br>/product="nonstructural protein"<br>/protein_id="AAB59851.1"<br>/db_xref="GI:333927"<br>/translation="MDTTHNDNTPQRLMTDRPLSLETITITSLTRDITTHKFIYILN<br>HBCIVRKLDKEQATFTFLVNYEMKLLKGVSTKYKKYTYNTKYCTFPMPFIINHDG<br>LFCIGIKPTKTPILYKVDLNP"<br>747<br>/gene="1B"<br>/note="a in [11]; g in [6]"<br>/replace="g"<br>769<br>variation |

```
/gene="IB"
/note="g in [11]; a in [6]"
/replace="a"
1081..2277
/gene="N"
1081..2277
/gene="N"
1096..2271
/gene="N"
/codon_start=1
/product="major nucleocapsid"
/protein_id="AAB59852.1"
/db_xref="GI:333928"
/translation="MALSKVLNDLTKDQLSSKVTIORSTGSDIDTPNDYVQKH
NKLGMILLITEDANHKFTGLIGMAYMSRLGREDTIKILRDAGYHVKANGVDVTHQ
DINGKMKFVLTLTASLTETIQINIEISRSKYKMLKEMGEVAPETFRSDPCGMII
LCIALVITKLAAGDRGLTAVIRRANNVLNEMRYKGLLPKDIAINSFYEVFEKPH
FIDVFVHFHGAQSVTRSGSRVEGIFAGLFMAYGAGQVMLFGLWGLAKSVKNIMLGHAS
VQAEQVEYEAQKLGGEAGFYHILNPKASLLSLTOPPHFSSVVLGNAAGLGM
GEYRGTPRNQDLYDAKAYARQLKENGVINYSVLDTAELEIAIKHQLNPKNDVEL"
2289..3196
/gene="P"
2289..3196
/gene="P"
2302..3027
/gene="P"
/codon_start=1
/product="phosphoprotein"
/protein_id="AAB59853.1"
/db_xref="GI:333929"
/translation="MEKFAPEFHGEDANNRATKFLSIKGFSPDKKKDSIIISVN
SIDIEVTKESPTISNTIINPTNEDDTAGNKPNYQRKPLVSKEDPTPSDNPFSKLY
KETIETPDNNEESSYSEIINDQTDNITARDRIDELKSELGLMLHTLVASAGPT
SARDGIRDAMIGUREEMIEKIRTEALMTNDRLEMARLRNESEKMAKDTSDVSLNP
```

Query Match 91.6%; Score 654.8; DB 14; Length 8510;  
Best local Similarity 94.8%; Pred. No. 1.1e-137;  
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
QY 1 CACAAGTCACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60
Db 4839 CACAAGTCACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 4898
QY 61 ACCCCACATACCTCACTCAGGATCCTCAGCTTGGATCAGTCTCCCAATCTGTCTGAA 120
Db 4899 ACCCCACATACCTCACTCAGGATCCTCAGCTTGGATCAGTCTCCCAATCTGTCTGAA 4958
QY 121 ATTACATCAAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAAACCTG 180
Db 4959 ATTACATCAAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAAACCTG 5018
QY 181 CAACCCACACAGTCAAGACTAATAACACACACACACACACACACACACACACACAC 240
Db 5019 CAATCCAACACAGTCAAGACTAATAACACACACACACACACACACACACACACAC 5078
QY 241 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5079 ACCACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5138
QY 301 GTGTTTAACTTTGTACCTCGAGCATATGACGACAACTCAACCTGCTGGGCTATCTGC 360
Db 5139 GTGTTTAACTTTGTACCTCGAGCATATGACGACAACTCAACCTGCTGGGCTATCTGC 5198
QY 361 AAAAGATACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
Db 5199 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5258
QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
Db 5259 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5318
QY 481 ACCACCAAGCCCAAGAGGACCAACCATCAACACCAACCAACCAACCAACCAACCA 540
```

```
Db 5319 ACCACCAAGCCCAAGAGGACCAACCATCAACACCAACCAACCAACCAACCAAC 5378
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
Db 5379 CTACTCACTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5438
QY 601 TCAACCTCTCTCCGAAAGCAATCTAAGCCCTTCTCAAGTCTCCCAACCAATCCGAGACCCA 660
Db 5439 TCAACTTCTCTCCGAAAGCAATCTAAGCCCTTCTCAAGTCTCTCAACATCCGAGTACCCA 5498
QY 661 TCACAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 714
Db 5499 TCACAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 5552
```

Search completed: October 30, 2003, 00:01:08  
Job time : 2924.85 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:32:25 ; Search time 54.717 Seconds  
(without alignments)  
1094.143 Million cell updates/sec

Title: US-09-462-816-4  
Perfect score: 1229  
Sequence: 1 HKVLTALIQDARSQIKNT.....VSTTSEHPSPQSSPPNTRQ 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL.23:\*  
1: sp archea:\*  
2: sp bacteria:\*  
3: sp fungi:\*  
4: sp human:\*  
5: sp invertebrate:\*  
6: sp mammal:\*  
7: sp mhc:\*  
8: sp organelle:\*  
9: sp phase:\*  
10: sp plant:\*  
11: sp rodent:\*  
12: sp virus:\*  
13: sp vertebrate:\*  
14: sp unclassified:\*  
15: sp rvirus:\*  
16: sp bacteriap:\*  
17: sp archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Match | Length | ID        | Description        |
|------------|-------|-------|--------|-----------|--------------------|
| 1          | 1229  | 100.0 | 298    | 12 Q82068 | Q82068 human respi |
| 2          | 1229  | 100.0 | 298    | 12 Q01929 | Q01929 human respi |
| 3          | 1033  | 88.9  | 298    | 12 Q09719 | Q09719 respiratory |
| 4          | 1093  | 88.9  | 298    | 12 Q09634 | Q09634 human respi |
| 5          | 1088  | 88.5  | 279    | 12 Q09678 | Q09678 human respi |
| 6          | 1072  | 87.2  | 279    | 12 Q096U0 | Q096U0 human respi |
| 7          | 1067  | 86.8  | 297    | 12 Q9YVB5 | Q9YVB5 human respi |
| 8          | 1066  | 86.7  | 297    | 12 Q82066 | Q82066 human respi |
| 9          | 1063  | 86.5  | 293    | 12 Q9YVB3 | Q9YVB3 human respi |
| 10         | 1062  | 86.4  | 298    | 12 Q9YVB4 | Q9YVB4 human respi |
| 11         | 1056  | 85.9  | 297    | 12 Q82071 | Q82071 human respi |
| 12         | 1055  | 85.8  | 278    | 12 Q09673 | Q09673 human respi |
| 13         | 1053  | 85.7  | 292    | 12 Q9YVB2 | Q9YVB2 human respi |
| 14         | 1053  | 85.7  | 297    | 12 Q82058 | Q82058 human respi |
| 15         | 1052  | 85.6  | 295    | 12 Q86359 | Q86359 respiratory |
| 16         | 1047  | 85.2  | 297    | 12 Q82067 | Q82067 human respi |

|    |      |      |     |           |                    |
|----|------|------|-----|-----------|--------------------|
| 17 | 1047 | 85.2 | 297 | 12 Q81947 | Q81947 human respi |
| 18 | 1045 | 85.0 | 278 | 12 Q906T1 | Q906T1 human respi |
| 19 | 1045 | 85.0 | 295 | 12 Q86360 | Q86360 respiratory |
| 20 | 1044 | 84.9 | 279 | 12 Q90680 | Q90680 human respi |
| 21 | 1044 | 84.9 | 279 | 12 Q906R9 | Q906R9 human respi |
| 22 | 1044 | 84.9 | 293 | 12 Q9YVB1 | Q9YVB1 human respi |
| 23 | 1043 | 84.9 | 298 | 12 Q82065 | Q82065 human respi |
| 24 | 1041 | 84.7 | 278 | 12 Q906U1 | Q906U1 human respi |
| 25 | 1040 | 84.6 | 297 | 12 Q82074 | Q82074 human respi |
| 26 | 1040 | 84.6 | 297 | 12 Q82064 | Q82064 human respi |
| 27 | 1039 | 84.5 | 279 | 12 Q906S2 | Q906S2 human respi |
| 28 | 1039 | 84.5 | 292 | 12 Q9YVB0 | Q9YVB0 human respi |
| 29 | 1037 | 84.4 | 279 | 12 Q906S9 | Q906S9 human respi |
| 30 | 1037 | 84.4 | 295 | 12 Q86361 | Q86361 respiratory |
| 31 | 1037 | 84.4 | 297 | 12 Q82063 | Q82063 human respi |
| 32 | 1036 | 84.3 | 278 | 12 Q906T4 | Q906T4 human respi |
| 33 | 1036 | 84.3 | 279 | 12 Q906S6 | Q906S6 human respi |
| 34 | 1035 | 84.2 | 279 | 12 Q906T9 | Q906T9 human respi |
| 35 | 1035 | 84.2 | 299 | 12 Q82077 | Q82077 human respi |
| 36 | 1033 | 84.1 | 279 | 12 Q906S7 | Q906S7 human respi |
| 37 | 1033 | 84.1 | 295 | 12 Q86357 | Q86357 respiratory |
| 38 | 1032 | 84.0 | 279 | 12 Q906S1 | Q906S1 human respi |
| 39 | 1032 | 84.0 | 297 | 12 Q9YVC8 | Q9YVC8 human respi |
| 40 | 1032 | 84.0 | 297 | 12 Q91946 | Q91946 human respi |
| 41 | 1031 | 83.9 | 298 | 12 Q82060 | Q82060 human respi |
| 42 | 1029 | 83.7 | 297 | 12 Q82072 | Q82072 human respi |
| 43 | 1028 | 83.6 | 297 | 12 Q82073 | Q82073 human respi |
| 44 | 1028 | 83.6 | 297 | 12 Q82057 | Q82057 human respi |
| 45 | 1028 | 83.6 | 298 | 12 Q82062 | Q82062 human respi |

#### ALIGNMENTS

#### RESULT 1

Q82068 ID Q82068 PRELIMINARY; PRT; 298 AA.  
AC Q82068;  
DT 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE Glycoprotein (Fragment).  
GN G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN= subgroup A;  
RX MEDLINE=94335057; PubMed=8057427;  
RA Garcia O., Martin M., Dopazo J., Arbiza J., Fabrisile S., Russi J.,  
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
RA Melero J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): cocirculating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein.";  
RL J. Virol. 68:5448-5459(1994).  
DR EMBL; Z33429; CAA83872.1;  
DR InterPro; IPR000925; Glycoprot G.  
DR InterPro; IPR006162; Ppantne attach.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.  
FT NON TER 1 1  
SQ SEQUENCE 298 AA; 32753 MW; 7D080307897A772B CRC64;

Query Match 100.0%; Score 1229; DB 12; Length 298;

Best Local Similarity 100.0%; Pred. No. 3.8e-89;

Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTALIQDARSQIKNTPTLTQDPQLGFSNLSITSTQTTILASTPGVKSNI 60  
|||||  
DB 67 HKVLTALIQDARSQIKNTPTLTQDPQLGFSNLSITSTQTTILASTPGVKSNI 126

```

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OY 61 OPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFEVFNFPVCSICSNPTCWAIC 120
DB 127 QPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFEVFNFPVCSICSNPTCWAIC 186
OY 121 KRIPNKKPGKTKTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTSEPTINTKINITT 180
DB 187 KRIPNKKPGKTKTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTSEPTINTKINITT 246
OY 181 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 232
DB 247 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 298

RESULT 2
OY1929 PRELIMINARY; PRT; 298 AA.
AC Q01929;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE G protein.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91065351; PubMed=2249671;
RA Garcia-Barreno B., Portela A., Delgado T., Lopez J.A., Melero J.A.;
RT "Frame shift mutations as a novel mechanism for the generation of
RT neutralization resistant mutants of human respiratory syncytial
RT virus.";
RL EMBO J. 9:4181-4187(1990).
DR EMBL; X17085; CAA34937.1; -
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne_attach.
DR Pfam; PF08002; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
SQ SEQUENCE 298 AA; 32779 MW; F8CD4213D97C2952 CRC64;

Query Match 100.0%; Score 1229; DB 12; Length 298;
Best Local Similarity 100.0%; Pred. No. 3.8e-89;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HKVLTLLAIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTQTTLASTTPGVKSNL 60
DB 67 HKVLTLLAIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTQTTLASTTPGVKSNL 126
OY 61 OPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFEVFNFPVCSICSNPTCWAIC 120
DB 127 QPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFEVFNFPVCSICSNPTCWAIC 186
OY 121 KRIPNKKPGKTKTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTSEPTINTKINITT 180
DB 187 KRIPNKKPGKTKTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTSEPTINTKINITT 246
OY 181 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 232
DB 247 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 298

RESULT 3
OY09719 PRELIMINARY; PRT; 298 AA.
AC O09719;
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Attachment glycoprotein (G).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

```

```

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OY NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2 ts1C;
RX MEDLINE=97185152; PubMed=9032893;
RA Tolley K.P., Marriott A.C., Simpson A., Plows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle C.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL Vaccine 14:1637-1646 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S2 ts1C;
RA Easton A.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39661; AAC57036.1; -
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne_attach.
DR Pfam; PF08002; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
SQ SEQUENCE 298 AA; 32779 MW; 67F4A043682FA450 CRC64;

Query Match 88.9%; Score 1093; DB 12; Length 298;
Best Local Similarity 90.1%; Pred. No. 1.9e-78;
Matches 209; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

OY 1 HKVLTLLAIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTQTTLASTTPGVKSNL 60
DB 67 HKVLTLLAIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTQTTLASTTPGVKSNL 126
OY 61 OPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFEVFNFPVCSICSNPTCWAIC 120
DB 127 QPTTVKTKNTTTTQTPSKPTTKORONKPNKPNNDHFEVFNFPVCSICSNPTCWAIC 186
OY 121 KRIPNKKPGKTKTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTSEPTINTKINITT 180
DB 187 KRIPNKKPGKTKTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTSEPTINTKINITT 246
OY 181 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 232
DB 247 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 298

RESULT 4
OY09634 PRELIMINARY; PRT; 298 AA.
AC O09634;
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Attachment protein (G).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RX MEDLINE=97185152; PubMed=9032893;
RA Tolley K.P., Marriott A.C., Simpson A., Plows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle C.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL Vaccine 14:1637-1646 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RA Easton A.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; U39662; AAC57026.1; -.  
DR InterPro; IPR000925; Glycoprot G.  
DR InterPro; IPR006162; Ppanthe attach.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.  
SQ SEQUENCE 298 AA; 32749 MW; 60F27B29D4997F31 CRC64;

Query Match 88.9%; Score 1093; DB 12; Length 298;  
Best Local Similarity 90.1%; Pred. No. 1.9e-78;  
Matches 209; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 1 HKVLTLLTALIQDTSQIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 60  
DB 67 HKVLTLLTALIQDTSQIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 126  
QY 61 QPTVTKNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTVTKNTTQIKQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTKTKPTKPTKPTKTKKDLKPQTTKPKVEPTTKTEPTINTKNTITTT 180  
DB 187 KRIPNKKPGKTKTKPTKPTKPTKTKKDLKPQTTKPKVEPTTKTEPTINTKNTITTT 246  
QY 181 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTTQ 232  
DB 247 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTTQ 298

## RESULT 5

Q9Q6T8 PRELIMINARY; PRT; 279 AA.

DT 01-MAY-2000 (TremBLrel. 13, Created)  
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
DE Glycoprotein (Fragment).

OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=92011;  
RX MEDLINE=20283719; PubMed=10823752;

RA Choi E.H., Lee H.J.;  
RT "Genetic diversity and molecular epidemiology of the G protein of subgroups A and B of respiratory syncytial virus isolated over 9 consecutive epidemics in Korea.";  
RL J. Infect. Dis. 181:1547-1556(2000).  
DR EMBL; AF193308; AAF23731.1; -.  
DR InterPro; IPR000925; Glycoprot\_G.  
DR InterPro; IPR006162; Ppanthe attach.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.  
FT NON TER 1  
SQ SEQUENCE 279 AA; 30498 MW; 4E1CDF2F7E569E3A CRC64;

Query Match 88.5%; Score 1088; DB 12; Length 279;  
Best Local Similarity 90.1%; Pred. No. 4.4e-78;  
Matches 209; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVLTLLTALIQDTSQIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 60  
DB 48 HKVLTLLTALIQDTSQIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 107  
QY 61 QPTVTKNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 108 LPTVTKNTTQIQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 167  
QY 121 KRIPNKKPGKTKTKPTKPTKPTKTKKDLKPQTTKPKVEPTTKTEPTINTKNTITTT 180  
DB 168 KRIPNKKPGKTKTKPTKPTKPTKTKKDLKPQTTKPKVEPTTKTEPTINTKNTITTT 227

QY 181 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTTQ 232  
DB 228 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTTQ 279

## RESULT 6

Q9Q6U0 PRELIMINARY; PRT; 279 AA.

AC Q9Q6U0;  
DT 01-MAY-2000 (TremBLrel. 13, Created)  
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
DE Glycoprotein (Fragment).  
GN G.

OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=91242;  
RX MEDLINE=20283719; PubMed=10823752;  
RA Choi E.H., Lee H.J.;  
RT "Genetic diversity and molecular epidemiology of the G protein of subgroups A and B of respiratory syncytial virus isolated over 9 consecutive epidemics in Korea.";  
RL J. Infect. Dis. 181:1547-1556(2000).  
DR EMBL; AF193306; AAF23729.1; -.  
DR InterPro; IPR000925; Glycoprot G.  
DR InterPro; IPR006162; Ppanthe attach.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.  
FT NON TER 1  
SQ SEQUENCE 279 AA; 30544 MW; 6B270AAEB1CB3533 CRC64;

Query Match 87.2%; Score 1072; DB 12; Length 279;  
Best Local Similarity 90.0%; Pred. No. 8e-77; Indels 0; Gaps 0;  
Matches 207; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVLTLLTALIQDTSQIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 60  
DB 48 HKVLTLLTALIQDTSQIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 107  
QY 61 QPTVTKNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 108 LPTVTKNTTQIQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 167  
QY 121 KRIPNKKPGKTKTKPTKPTKPTKTKKDLKPQTTKPKVEPTTKTEPTINTKNTITTT 180  
DB 168 KRIPNKKPGKTKTKPTKPTKPTKTKKDLKPQTTKPKVEPTTKTEPTINTKNTITTT 227  
QY 181 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTT 230  
DB 228 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTT 277

## RESULT 7

Q9YVB5 PRELIMINARY; PRT; 297 AA.

AC Q9YVB5;  
DT 01-MAY-1999 (TremBLrel. 10, Created)  
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)  
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
DE Attachment glycoprotein G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WV2780;  
RX MEDLINE=99022964; PubMed=9806017;



```

OX NCBI_TaxID=11250;
RN SEQUENCE FROM N.A.
RP SEQUENCE=99022964; PubMed=9806017;
RC STRAIN=99022964;
RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
RX "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932 (1998).
DR EMBL; AF065406; AAD02942.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
SQ SEQUENCE 298 AA; 32781 MW; AE7C9E05547C5745 CRC64;

Query Match 86.4%; Score 1062; DB 12; Length 298;
Best Local Similarity 87.5%; Pred. No. 5.2e-76;
Matches 203; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITTSQTTTILASTTPGVKSNL 60
DB 1 HKVTLTALIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITTSQTTTILASTTPGVKSNL 126
QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120
DB 127 QSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186
QY 121 KRIPNKKPGKTTTKPTKPTTKKDLKQPTTKPKVPTTKTEPTINTTKNITTT 180
DB 187 KRIPNKKPGKTTTKPTKPTTKKDLKQPTTKPKVPTTKTEPTINTTKNITTT 246
QY 181 LLTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPSPNTTR 231
DB 247 PLTSNTRNPELTSQMETHTSSGNLSPSQVSTTSBHPSPSPSPNTSR 297

RESULT 11
Q82071 ID Q82071 PRELIMINARY; PRT; 297 AA.
AC Q82071;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein.
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=94335057; PubMed=8057427;
RX Garcia O., Martin M., Dopazo J., Arbiza J., Fabraile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): co-circulating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459 (1994).
DR EMBL; Z33432; CAA83875.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
SQ SEQUENCE 297 AA; 32581 MW; AD947B2318193306 CRC64;

Query Match 85.9%; Score 1056; DB 12; Length 297;
Best Local Similarity 87.4%; Pred. No. 1.5e-75;
Matches 202; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITTSQTTTILASTTPGVKSNL 60
DB 1 HKVTLTALIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITTSQTTTILASTTPGVKSNL 126
QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120
DB 127 QSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186
QY 121 KRIPNKKPGKTTTKPTKPTTKKDLKQPTTKPKVPTTKTEPTINTTKNITTT 180
DB 187 KRIPNKKPGKTTTKPTKPTTKKDLKQPTTKPKVPTTKTEPTINTTKNITTT 246
QY 181 LLTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPSPNTTR 231
DB 247 PLTSNTRNPELTSQMETHTSSGNLSPSQVSTTSBHPSPSPSPNTSR 297

RESULT 12
Q906T3 ID Q906T3 PRELIMINARY; PRT; 278 AA.
AC Q906T3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=93006;
RX MEDLINE=20283719; PubMed=10823752;
RA Choi E.H., Lee H.J.;
RT "Genetic diversity and molecular epidemiology of the G protein of
RT subgroups A and B of respiratory syncytial virus isolated over 9
RT consecutive epidemics in Korea.";
RL J. Infect. Dis. 181:1547-1556 (2000).
DR EMBL; AF193313; AAF23736.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
FT NON TER 1
SQ SEQUENCE 278 AA; 30443 MW; 8E2374F2E1AE15C9 CRC64;

Query Match 85.8%; Score 1055; DB 12; Length 278;
Best Local Similarity 87.4%; Pred. No. 1.7e-75;
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITTSQTTTILASTTPGVKSNL 60
DB 48 HKVTLTALIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITTSQTTTILASTTPGVKSNL 107
QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120
DB 108 QSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 167
QY 121 KRIPNKKPGKTTTKPTKPTTKKDLKQPTTKPKVPTTKTEPTINTTKNITTT 180
DB 168 KRIPNKKPGKTTTKPTKPTTKKDLKQPTTKPKVPTTKTEPTINTTKNITTT 227
QY 181 LLTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPSPNTTR 230
DB 228 LLTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPSPNTTR 277

RESULT 13
Q9YVB2 ID Q9YVB2 PRELIMINARY; PRT; 292 AA.
AC Q9YVB2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)

```



```

DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Attachment glycoprotein G (Fragment).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WV19983;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932(1998).
DR EMBL; AF065408; AA02944.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; IPR002965; P rich extensn.
DR Pfam; PF08082; Glycoprotein_G; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR PROSITE; PS00012; PHOSPHOPANTHINE; 1.
FT NON_TER
FT SEQUENCE 292 AA; 31964 MW; 8942A8DD0A402A4B CRC64;

Query Match 85.7%; Score 1053; DB 12; Length 292;
Best Local Similarity 87.9%; Pred. No. 2.6e-75;
Matches 203; Conservative 8; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTPGVKSNL 60
DB 62 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTPGVKSTL 121
QY 61 OPTTVKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNFPVPCISCSNNPTCWAIC 120
DB 122 QSTTVGKNTTTTQAPNKPTTKQKQKPNKPNNDFFHFEVFNFPVPCISCSNNPTCWAIC 181
QY 121 KRIPNKKPGKKTTPKTKPTTKTKKDKLPQTKKQKQKPNNDFFHFEVFNFPVPCISCSNNPTT 180
DB 182 KRIPNKKPGKKTTPKTKPTTKTKKDKLPQTKKQKQKPNNDFFHFEVFNFPVPCISCSNNPTT 241
QY 181 LLTNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSPNTT 231
DB 242 LLTNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSPNTT 292

RESULT 14
Q82058 PRELIMINARY; PRT; 297 AA.
AC Q82058;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE (MAD-1-89) subgroup A, G glycoprotein.
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrisile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459(1994).
DR EMBL; Z33456; CAA83879.1; -.
DR InterPro; IPR000925; Glycoprot_G.

```

```

DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF08082; Glycoprotein_G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHINE; 1.
SQ SEQUENCE 297 AA; 32741 MW; 0E567A174BF64964 CRC64;

Query Match 85.7%; Score 1053; DB 12; Length 297;
Best Local Similarity 87.4%; Pred. No. 2.7e-75;
Matches 201; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTPGVKSNL 60
DB 67 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTPSAESTP 126
QY 61 OPTTVKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNFPVPCISCSNNPTCWAIC 120
DB 127 QSTTVKNTTTTQIQPSKPTTKQKQKPNKPNNDFFHFEVFNFPVPCISCSNNPTCWAIC 186
QY 121 KRIPNKKPGKKTTPKTKPTTKTKKDKLPQTKKQKQKPNNDFFHFEVFNFPVPCISCSNNPTT 180
DB 187 KRIPNKKPGKKTTPKTKPTTKTKKDKLPQTKKQKQKPNNDFFHFEVFNFPVPCISCSNNPTT 246
QY 181 LLTNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSPNTT 230
DB 247 LLTNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSPNTT 296

RESULT 15
Q86359 PRELIMINARY; PRT; 295 AA.
AC Q86359;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE G protein (fragment).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RA Cane P.A.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X73353; CAA51764.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF08082; Glycoprotein_G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHINE; 1.
FT NON_TER
FT SEQUENCE 295 AA; 32490 MW; 48D835F670FF8006 CRC64;

Query Match 85.6%; Score 1052; DB 12; Length 295;
Best Local Similarity 87.4%; Pred. No. 3.2e-75;
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTQTTILASTTPGVKSNL 60
DB 65 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITSTQTTTPATTPSAESTP 124
QY 61 OPTTVKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNFPVPCISCSNNPTCWAIC 120
DB 125 QSTTVKNTTTTQIQPSKPTTKQKQKPNKPNNDFFHFEVFNFPVPCISCSNNPTCWAIC 184
QY 121 KRIPNKKPGKKTTPKTKPTTKTKKDKLPQTKKQKQKPNNDFFHFEVFNFPVPCISCSNNPTT 180

```

Db 185 KRIENKKGKTTTKPTIKTKKDLKPQTTKPKVLTTRKTEKPTINTTRTNIRTT 244  
Oy 181 LLTNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQPPSPNTT 230  
Db 245 LLTNTTGNPEYTSQKETLHSTSEGNPSQVYTTSEYPSQPPSPNTT 294

Search completed: October 29, 2003, 17:40:53  
Job time : 55.717 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:12:25 ; Search time 12.6943 Seconds  
(without alignments)  
859.454 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVTLTTAIIDQASQIKNT.....VSTTSEHPSQSPSPNTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|-------|-------------|
| 1          | 1216  | 98.9  | 298    | 1     | VGLG_HRSVL  |
| 2          | 1120  | 91.1  | 298    | 1     | VGLG_HRSVA  |
| 3          | 1052  | 85.6  | 297    | 1     | VGLG_HRSV6  |
| 4          | 1045  | 85.0  | 297    | 1     | VGLG_HRSV3  |
| 5          | 1037  | 84.4  | 297    | 1     | VGLG_HRSV7  |
| 6          | 1033  | 84.1  | 297    | 1     | VGLG_HRSV4  |
| 7          | 1019  | 82.9  | 298    | 1     | VGLG_HRSV5  |
| 8          | 1009  | 82.1  | 297    | 1     | VGLG_HRSV2  |
| 9          | 505.5 | 41.1  | 292    | 1     | VGLG_HRSV8  |
| 10         | 497   | 40.4  | 292    | 1     | VGLG_HRSV1  |
| 11         | 202   | 16.4  | 307    | 1     | SGS3_DROME  |
| 12         | 198.5 | 16.2  | 2812   | 1     | ZAN_HUMAN   |
| 13         | 193   | 15.7  | 5179   | 1     | MUC2_HUMAN  |
| 14         | 191   | 15.5  | 662    | 1     | MUC1_XENLA  |
| 15         | 183   | 14.9  | 217    | 1     | SGS3_DROST  |
| 16         | 183   | 14.9  | 1161   | 1     | DANA_YEAST  |
| 17         | 180.5 | 14.7  | 263    | 1     | VGLG_BRST1  |
| 18         | 179.5 | 14.6  | 263    | 1     | VGLG_BRST4  |
| 19         | 177   | 14.4  | 257    | 1     | VGLG_BRSTC  |
| 20         | 176.5 | 14.4  | 263    | 1     | VGLG_BRSTW  |
| 21         | 170.5 | 13.9  | 263    | 1     | VGLG_BRST2  |
| 22         | 169.5 | 13.8  | 1367   | 1     | AMTH_YEAST  |
| 23         | 169   | 13.8  | 257    | 1     | VGLG_BRST7  |
| 24         | 166   | 13.5  | 2282   | 1     | ZAN_RABIT   |
| 25         | 164.5 | 13.4  | 2476   | 1     | ZAN_PIG     |
| 26         | 160   | 13.0  | 263    | 1     | SGS3_DROVA  |
| 27         | 159.5 | 13.0  | 263    | 1     | VGLG_ORSVW  |
| 28         | 159   | 12.9  | 257    | 1     | VGLG_BRST5  |
| 29         | 156   | 12.7  | 257    | 1     | VGLG_BRSTL  |
| 30         | 154   | 12.5  | 257    | 1     | VGLG_BRSTN  |
| 31         | 154   | 12.5  | 1233   | 1     | MUSA_HUMAN  |
| 32         | 153.5 | 12.5  | 5703   | 1     | MUSE_HUMAN  |
| 33         | 150.5 | 12.2  | 1513   | 1     | MUC2_RAT    |

## RESULT 1

VGLG\_HRSVL  
ID VGLG\_HRSVL STANDARD; PRT; 298 AA.  
AC P20895;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (subgroup A / strain Long).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11260;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=8728957; PubMed=2441388;  
RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;  
RT "The G glycoprotein of human respiratory syncytial viruses of  
RT subgroups A and B: extensive sequence divergence between  
RT antigenically related proteins";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).  
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.

-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M17212; AAA47411.1; -  
PIR; A32703; MGNZRL.  
InterPro; IPR000925; Glycoprot G.  
Pfam; PF00802; Glycoprotein G; 1.  
Transmembrane; Glycoprotein.  
DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66  
FT DOMAIN 67 298  
FT CARBOHYD 103 103 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 298 AA; 32781 MW; B79FEFA4BA73B0E CRC64;

Query Match 98.9%; Score 1216; DB 1; Length 298;  
Best Local Similarity 99.1%; Pred. No. 4.3e-75;  
Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M11486; AAB59857.1; -  
 DR EMBL; X03149; CAA26928.1; -  
 DR EMBL; U50362; AAB86663.1; -  
 DR EMBL; U50363; AAB86675.1; -  
 DR EMBL; U63644; AAC55969.1; -  
 DR PIR; A94048; MGZ;  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein\_G; 1.  
 DR Transmembrane; Glycoprotein.  
 KW DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66  
 FT DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 298 AA; 32586 MW; 993C3D2DD68BC634 CRC64;  
 Query Match 91.1%; Score 1120; DB 1; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.2e-68;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITSTQTTILASTTPGVKSNL 126  
 QY 61 OPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDFFHEFVFNVPVCSICSNNPTCWAIC 120  
 DB 127 OPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDFFHEFVFNVPVCSICSNNPTCWAIC 186  
 QY 121 KRIPNKKPKGKTKTKPTKPTKTKTKKOLKPKQTTKPKVEVPTTKPTERTPTNTTKNITTT 180  
 DB 187 KRIPNKKPKGKTKTKPTKPTKTKTKKOLKPKQTTKPKVEVPTTKPTERTPTNTTKNITTT 246  
 QY 181 LLTNNTTGNPKLTSQMETFSTSEGNLSQSVSTTSEHPSQSPSPNTTQ 232  
 DB 247 LLTNNTTGNPKLTSQMETFSTSEGNLSQSVSTTSEHPSQSPSPNTTQ 298  
 RESULT 2  
 VGLG\_HRSVA STANDARD; PRT; 298 AA.  
 AC P03423;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (strain A2).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11259;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85216636; PubMed=3858865;  
 RA Wertz G.W., Collins P.L., Huang Y., Gruber C., Levine S., Ball L.A.;  
 RT "Nucleotide sequence of the G protein gene of human respiratory  
 RT syncytial virus reveals an unusual type of viral membrane protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4075-4079(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86067198; PubMed=4069997;  
 RA Satake M., Colligan J.E., Elango N., Norrby E., Venkatesan S.;  
 RT "Respiratory syncytial virus envelope glycoprotein (G) has a novel  
 RT structure.";  
 RL Nucleic Acids Res. 13:7795-7812(1985).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95266253; PubMed=7747420;  
 RA Connors M., Crowe J.E. Jr., Firestone C.Y., Murphy B.R., Collins P.L.;  
 RT "A cold-passaged, attenuated strain of human respiratory syncytial  
 RT virus contains mutations in the F and L genes.";  
 RL Virology 208:478-484(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97187925; PubMed=9035372;  
 RA Crowe J.E. Jr., Firestone C.Y., Whitehead S.S., Collins P.L.,  
 RA Murphy B.R.;  
 RT "Acquisition of the ts phenotype by a chemically mutagenized cold-  
 RT passaged human respiratory syncytial virus vaccine candidate results  
 RT from the acquisition of a single mutation in the polymerase (L)  
 RT gene.";  
 RL Virus Genes 13:269-273(1996).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M11486; AAB59857.1; -  
 DR EMBL; X03149; CAA26928.1; -  
 DR EMBL; U50362; AAB86663.1; -  
 DR EMBL; U50363; AAB86675.1; -  
 DR EMBL; U63644; AAC55969.1; -  
 DR PIR; A94048; MGZ;  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein\_G; 1.  
 DR Transmembrane; Glycoprotein.  
 KW DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66  
 FT DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 298 AA; 32586 MW; 993C3D2DD68BC634 CRC64;  
 Query Match 91.1%; Score 1120; DB 1; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.2e-68;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDATSQIKNTTPTLTQDPQLGIFSLSNLSSEITSTQTTILASTTPGVKSTL 126  
 QY 61 OPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDFFHEFVFNVPVCSICSNNPTCWAIC 120  
 DB 127 OPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDFFHEFVFNVPVCSICSNNPTCWAIC 186  
 QY 121 KRIPNKKPKGKTKTKPTKPTKTKTKKOLKPKQTTKPKVEVPTTKPTERTPTNTTKNITTT 180  
 DB 187 KRIPNKKPKGKTKTKPTKPTKTKTKKOLKPKQTTKPKVEVPTTKPTERTPTNTTKNITTT 246  
 QY 181 LLTNNTTGNPKLTSQMETFSTSEGNLSQSVSTTSEHPSQSPSPNTTQ 232  
 DB 247 LLTNNTTGNPKLTSQMETFSTSEGNLSQSVSTTSEHPSQSPSPNTTQ 298  
 RESULT 3  
 VGLG\_HRSV6 STANDARD; PRT; 297 AA.  
 AC P27025;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (strain r6b256)  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11256;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91374005; PubMed=1895054;  
 RA Cane P.A., Matthews D.A., Pringle C.R.;  
 RT "Identification of variable domains of the attachment (G) protein of  
 RT subgroup A respiratory syncytial viruses.";  
 RL J. Gen. Virol. 72:2091-2096(1991).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC -----  
 CC PIR; J01208; J01208.  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein\_G; 1.

Transmembrane; Glycoprotein.

Query Match 85.6%; Score 1052; DB 1; Length 297;  
Best Local Similarity 87.4%; Pred. No. 4.2e-64;  
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVLTUTTAIIQDATSQIKNTTPTVLTQDPOLGISFNLSEISQTTTIIASTPGVKSNL 60  
DB 67 HKVLTUTTAIIQDATSQIKNTTPTVLTQDPOLGISFNLSEISQTTTPTAPTPSAESTP 126  
QY 61 QPTTVTKNTTUTTOTQSKPTTKORQKNPKNNDFEVENFVPCSCSNPTCWAIC 120  
DB 127 QSTTVTKNTTUTTQIQSPKPTTKORQKNPKNNDFEVENFVPCSCSNPTCWAIC 186  
QY 121 KRIPNKPQKTKTTPKTKPTTKKDLKQPTTKPEVPTTKTBEPTINTKNTTT 180  
DB 187 KRIPNKPQKTKTTPKTKPTTKKDLKQPTTKPEVPTTKTBEPTINTKNTTT 246  
QY 181 LTTNNTNGPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPQSPSPNTT 230  
DB 247 LTTNNTNGPEYTSQKFLHSTSPGNPSPSOVTTTSEYPPQSPSPNTT 296

RESULT 4  
VGLG HRSV3  
ID VGLG HRSV3 STANDARD; PRT; 297 AA.  
AC P27022;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GS G.  
OS Human respiratory syncytial virus (strain rsb1734).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OC NCBI\_TaxId=11253;  
RN [1]  
RS SEQUENCE FROM N.A.  
RX MEDLINE=91374005; Pubmed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses.";  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
CC FIR; JQ1205; JQ1205.  
DR InterPro; IPR000925; Glycoprot G.  
DR Pfam; PF00802; Glycoprotein G; 1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 103 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 297 AA; 32525 MW; 6781756C38B64A80 CRC64;

Query Match 85.0%; Score 1045; DB 1; Length 297;  
Best Local Similarity 86.6%; Pred. No. 1.3e-63;

QY 1 HKVLTUTTAIIQDATSQIKNTTPTVLTQDPOLGISFNLSEISQTTTIIASTPGVKSNL 60  
DB 67 HKVLTUTTAIIQDATSQIKNTTPTVLTQDPOLGISFNLSEISQTTTPTAPTPSAESTP 126  
QY 61 QPTTVTKNTTUTTOTQSKPTTKORQKNPKNNDFEVENFVPCSCSNPTCWAIC 120  
DB 127 QSTTVTKNTTUTTQIQSPKPTTKORQKNPKNNDFEVENFVPCSCSNPTCWAIC 186  
QY 121 KRIPNKPQKTKTTPKTKPTTKKDLKQPTTKPEVPTTKTBEPTINTKNTTT 180  
DB 187 KRIPNKPQKTKTTPKTKPTTKKDLKQPTTKPEVPTTKTBEPTINTKNTTT 246  
QY 181 LTTNNTNGPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPQSPSPNTT 230  
DB 247 LTTNNTNGPEYTSQKFLHSTSPGNPSPSOVTTTSEYPPQSPSPNTT 296

RESULT 4  
VGLG HRSV3  
ID VGLG HRSV3 STANDARD; PRT; 297 AA.  
AC P27022;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GS G.  
OS Human respiratory syncytial virus (strain rsb1734).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OC NCBI\_TaxId=11253;  
RN [1]  
RS SEQUENCE FROM N.A.  
RX MEDLINE=91374005; Pubmed=1895054;  
RA Cane P.A., Matthews D.A., Pringle C.R.;  
RT Identification of variable domains of the attachment (G) protein of  
RT subgroup A respiratory syncytial viruses.";  
RL J. Gen. Virol. 72:2091-2096(1991).  
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
CC FIR; JQ1205; JQ1205.  
DR InterPro; IPR000925; Glycoprot G.  
DR Pfam; PF00802; Glycoprotein G; 1.  
KW Transmembrane; Glycoprotein.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 297 AA; 32525 MW; 48443F9E091E1802 CRC64;



RN SEQUENCE FROM N.A.  
 RP MEDLINE=91374005; PubMed=1895054;  
 RA Cane P.A., Matthews D.A., Pringle C.R.;  
 RT "Identification of variable domains of the attachment (G) protein of  
 RT subgroup A respiratory syncytial viruses."  
 RL J. Gen. Virol. 72:2091-2096(1991).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR PIR; JQ1204; JQ1204.  
 DR InterPro: IPR000925; Glycoprot G.  
 DR Pfam: PF00802; Glycoprotein\_G\_1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66 POTENTIAL.  
 FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 297 AA; 32745 MW; FC72A7F3A8EF67C CRC64;  
  
 Query Match 82.1%; Score 1009; DB 1; Length 297;  
 Best Local Similarity 84.0%; Pred. No. 3.2e-61;  
 Matches 194; Conservative 11; Mismatches 26; Indels 0; Gaps 0;  
  
 QY 1 HKVLTITAIIDATSOIKNTTPTLTQDPQIGISFNSLSEITSTQTTILASTTGVKSNL 60  
 DB 67 HRVSTSTIIQDATNQIKNTTPTLTQDPQIGISFNSLSEITSLITLIDPTTGVKLT 126  
  
 QY 61 QPTTVKTKNTTITQPSKPTTKQKPNKPNNDHFVFNVPVCSICSNPTCWAIC 120  
 DB 127 QSTTVIKNTTITQAPKSKTKQKPNKPNNDHFVFNVPVCSICSNPTCWAIC 186  
  
 QY 121 KRIPNKKPKGKTTTKTKPKTKTKKDLKPKQTKPKKEVPTTKPTTEPTINTKNTTT 180  
 DB 187 KRIPNKKPKGKTTTKTKPKTKTKKDLKPKQTKPKKEVPTTKPTTEPTINTKNTTT 246  
  
 QY 181 LLTNTNTGNPKLTQSMETFTSTSEGNLSPSOVSTTSEHPSQSPPTNR 231  
 DB 247 PLTSNARNPLTSQMETFTSTSEGNLSPSOVSTTSEHPSQSPPTNR 297  
  
 RESULT 9  
 VGLG HRSV8  
 ID VGLG HRSV8 STANDARD; PRT; 292 AA.  
 AC P23041.  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (subgroup B / strain 8/60).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11258;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=90357765; PubMed=1697126;  
 RA Sullender W.M., Anderson K., Wertz G.W.;  
 RT "The respiratory syncytial virus subgroup B attachment glycoprotein:  
 RT analysis of sequence, expression from a recombinant vector, and  
 RT evaluation as an immunogen against homologous and heterologous  
 RT subgroup virus challenge."  
 RL Virology 178:195-203(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=91374595; PubMed=1895391;  
 RX Sullender W.M., Mufson M.M., Anderson L.J., Wertz G.W.;

RT "Genetic diversity of the attachment protein of subgroup B  
 RT respiratory syncytial viruses."  
 RL J. Virol. 65:5425-5434(1991).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; M55633; AAA47413.1; .  
 DR EMBL; M73545; AAA47408.1; .  
 DR PIR; A37077; MGNZ60.  
 DR InterPro: IPR000925; Glycoprot G.  
 DR Pfam: PF00802; Glycoprotein\_G\_1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66 POTENTIAL.  
 FT DOMAIN 67 292 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 292 AA; 32143 MW; 8EC6C85EF057BBS CRC64;  
  
 Query Match 41.1%; Score 505.5; DB 1; Length 292;  
 Best Local Similarity 47.6%; Pred. No. 1.9e-27;  
 Matches 108; Conservative 28; Mismatches 88; Indels 3; Gaps 3;  
  
 QY 1 HKVLTITAIIDATSOIKNTTPTLTQDPQIGISFNSLSEITSTQTTILASTTGVKSNL 60  
 DB 67 HKVLTITVTVQTKNTGKNTSTLTQVPPERVNSKKQPTTSTPIHNSATISPNTKSET 126  
  
 QY 61 QPTTVKTKNTTITQPSKPTTKQKPNKPNNDHFVFNVPVCSICSNPTCWAIC 120  
 DB 127 HHTTAQTKGRITTSQTKNPKSKSNPKPKDDHFEVFNVPVCSICGNQLCKSIC 186  
  
 QY 121 KRIPNKKPKGKTTTKTKPKTKTKKDLKPKQTKPKKEVPTTKPTTEPTINTKNTTT 179  
 DB 187 KRIPNKKPKGKTTTKTKPKTKTKKDLKPKQTKPKKEVPTTKPTTEPTINTKNTTT 246  
  
 QY 180 TLLTNTNTGNPKLTQSMETFTSTSEGNLSPSOVSTTSEHPSQSPSP 226  
 DB 247 SQSTVIDITTPKYTIQQSLHSTSTSEGNLSPSOVSTTSEHPSQSPSP 291  
  
 RESULT 10  
 VGLG HRSV1  
 ID VGLG HRSV1 STANDARD; PRT; 292 AA.  
 AC P20896;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-NOV-1991 (Rel. 20, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (subgroup B / strain 18537).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11251;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87289657; PubMed=2441388;  
 RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;  
 RT "The G glycoprotein of human respiratory syncytial viruses of  
 RT subgroups A and B: extensive sequence divergence between



antigenically related proteins.";

Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629 (1987).

-1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND HEMAGGLUTININATING ACTIVITIES.

-1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.

-1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; M17213; AAA47412.1; --

PIR; B32703; MGZ18.

InterPro; IPR000925; Glycoprot G.

Pfam; PF00802; Glycoprotein\_G; 1.

Transmembrane; Glycoprotein.

DOMAIN 1 37

TRANSMEM 38 66

DOMAIN 67 292

FT CARBOHYD 81 81

FT CARBOHYD 86 86

FT CARBOHYD 100 100

SEQUENCE 292 AA; 32306 MW; BC8C59F69CA7AFC2 CRC64;

Query Match 40.48; Score 497; DB 1; Length 292;

Best Local Similarity 46.88; Pred. No. 6.9e-27;

Matches 104; Conservative 28; Mismatches 88; Indels 2; Gaps 2;

QY 1 HKVTLTAIQDAPTSQIKNTPTTYLTQDPQLGIFSLSLSEITSTTILASTTGVKSNL 60

DB 67 HKVTLTVTQTIKHTKNTSTLTQVPERVNSKQPTTSPHTNSATISNTWSET 126

QY 61 QPTVTKNTTITQTSKPTTKORQKNPKNDHFVFNVPVCSISNPTCAIC 120

DB 127 HHTTAQTKGRITSTQTNKESTKSRNPKPKKDDYHFEVFNVPVCSICGNQLCKSK 186

QY 121 KRIPNKPKGKTTTKPKTKPTKT--KKDLKPTQTKPEVPTTKPTBEPTINTKNTIT 179

DB 187 KTIPTNPKPKTKPTK 246

QY 180 TLTNTNTGNPKTSQMTSTHSISSEGNLSPSQVSTTSEHPS 221

DB 247 SQSTVLDITPKVTIQQSLHSTSTENTPSQTPTASE-PS 287

RESULT 11

SGS3 DROME STANDARD; PRT; 307 AA.

AC P02840; Q9VJ2;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Salivary glue protein Sgs-3 precursor.

GN Sgs3 OR CG11720.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83294545; PubMed=6411930;

RA Garfinkel M.D., Pruitt R.B., Meyerowitz E.M.;

RT "DNA sequences, gene regulation and modular protein evolution in the Drosophila 68C glue gene cluster.";

J. Mol. Biol. 168:765-789 (1983).

RN RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Liakou P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svendsen R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

Science 287:2185-2195 (2000).

[3]

RN RP SEQUENCE OF 1-28 FROM N.A.

RC MEDLINE=88332966; PubMed=3138416;

RA Martin C.H., Mayeda C.A., Meyerowitz E.M.;

RT "Evolution and expression of the Sgs-3 glue gene of Drosophila.";

J. Mol. Biol. 201:273-287 (1988).

[4]

RN RP DEVELOPMENTAL STAGE.

RX MEDLINE=94038699; PubMed=8223281;

RA Huet F., Ruiz C., Richards G.;

RT "Puffs and PCR: the in vivo dynamics of early gene expression during eclosion responses in Drosophila.";

Development 118:613-627 (1993).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: SALIVARY GLAND SPECIFIC.

CC -1- DEVELOPMENTAL STAGE: IN THE SALIVARY GLANDS OF MID INSTAR LARVAE LEVELS DRAMATICALLY INCREASE DURING PUFF STAGE 1 AT 98-106 HOURS OF DEVELOPMENT. LEVELS REMAIN CONSTANT AND ABUNDANT IN LATE LARVAE UNTIL PUFF STAGE 10, THEN DECREASE BY STAGE 11.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X01918; CAA25994.1; --

DR EMBL; AE003544; BAF50056.1; --

DR EMBL; X78392; CAA55154.1; --





"MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymorphism";  
 RL J. Clin. Invest. 88:1005-1013(1991).  
 CC -|- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND  
 CC OTHER MUCUS MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A  
 CC PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIOUS  
 CC AGENTS AT MUCOSAL SURFACES.  
 CC -|- SUBUNIT: MULTIMERIC.  
 CC -|- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,  
 CC BRONCHUS, CERVIX AND GALL BLADDER.  
 CC -|- FTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR  
 CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).  
 CC -|- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND  
 CC VARIES AMONG DIFFERENT ALLELES.  
 CC -|- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.  
 CC -|- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.  
 CC -|- SIMILARITY: Contains 2 WFEC domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; L21998; AAB95295.1; -  
 DR EMBL; M74027; AAA59875.1; -  
 DR EMBL; M94131; AAA59163.1; -  
 DR EMBL; M94132; AAA59164.1; -  
 DR PIR; A49963; A43932.  
 DR Genew; HGNC:7512; MUC2.  
 DR MIM; 158370; -  
 DR GO; GO:0005803; C:secretory vesicle; TAS.  
 DR InterPro; IPR006208; Cys knot.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR002919; TIL\_Cysrich.  
 DR InterPro; IPR001007; VWF\_C.  
 DR InterPro; IPR001846; VWF\_D.  
 DR Pfam; PF00007; Cys knot; 1.  
 DR Pfam; PF00093; vwc; 1.  
 DR Pfam; PF00094; vwd; 4.  
 DR SMART; SM00214; VWC; 2.  
 DR SMART; SM00216; VWD; 4.  
 DR PROSITE; PS01185; CTCK 1; 1.  
 DR PROSITE; PS01225; CTCK 2; 1.  
 DR PROSITE; PS00022; EGF 1; UNKNOWN 1.  
 DR PROSITE; PS01208; VWF 1; 2.  
 DR PROSITE; PS0184; VWF 2; 2.  
 KW Glycoprotein; Repeat; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 5179 MUCIN 2.  
 FT DOMAIN 1401 1747 APPROXIMATE REPEATS.  
 FT REPEAT 1401 1416 1.  
 FT REPEAT 1417 1432 2.  
 FT REPEAT 1433 1448 3.  
 FT REPEAT 1449 1464 4.  
 FT REPEAT 1465 1471 5.  
 FT REPEAT 1472 1478 6.  
 FT REPEAT 1479 1494 7A.  
 FT REPEAT 1495 1517 7B.  
 FT REPEAT 1518 1533 8A.  
 FT REPEAT 1534 1556 8B.  
 FT REPEAT 1557 1572 9A.  
 FT REPEAT 1573 1596 9B.  
 FT REPEAT 1597 1612 10A.  
 FT REPEAT 1613 1635 10B.  
 FT REPEAT 1636 1651 11A.  
 FT REPEAT 1652 1675 11B.  
 FT REPEAT 1676 1683 12.

FT REPEAT 1684 1699  
 FT REPEAT 1700 1715  
 FT REPEAT 1716 1731  
 FT REPEAT 1732 1747  
 FT DOMAIN 4815 4886  
 FT DOMAIN 4924 4991  
 FT DOMAIN 5075 5160  
 FT DISULFID 5089 5136  
 FT DISULFID 5098 5152  
 FT DISULFID 5102 5154  
 FT DISULFID ? 5159  
 FT CARBOHYD 163 183  
 FT CARBOHYD 423 423  
 FT CARBOHYD 670 670  
 FT CARBOHYD 770 770  
 FT CARBOHYD 894 894  
 FT CARBOHYD 1139 1139  
 FT CARBOHYD 1154 1154  
 FT CARBOHYD 1215 1215  
 FT CARBOHYD 1230 1230  
 FT CARBOHYD 1246 1246  
 FT CARBOHYD 1787 1787  
 FT CARBOHYD 1820 1820  
 FT CARBOHYD 4339 4339  
 FT CARBOHYD 4351 4351  
 FT CARBOHYD 4362 4362  
 FT CARBOHYD 4373 4373  
 FT CARBOHYD 4422 4422  
 FT CARBOHYD 4438 4438  
 FT CARBOHYD 4502 4502  
 FT CARBOHYD 4616 4616  
 FT CARBOHYD 4627 4627  
 FT CARBOHYD 4752 4752  
 FT CARBOHYD 4787 4787  
 FT CARBOHYD 4881 4881  
 FT CARBOHYD 4888 4888  
 FT CARBOHYD 4955 4955  
 FT CARBOHYD 4970 4970  
 FT CARBOHYD 5019 5019  
 FT CARBOHYD 5038 5038  
 FT CARBOHYD 5069 5069  
 FT CONFLICT 1351 1351  
 FT CONFLICT 1412 1412  
 FT CONFLICT 1449 1449  
 FT CONFLICT 1504 1504  
 FT CONFLICT 4192 4192  
 SQ SEQUENCE 5179 AA; 540295 MM; 85CD7571FB9A5663 CRC64;  
 Query Match 15.7%; Score 193; DB 1; Length 5179;  
 Best Local Similarity 30.6%; Pred. No. 3.3e-05;  
 Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;  
 QY 5 LTTAIIQDA--TSQIKNTTP-----TYLTQDQLGISFSNLSEIT--SQTTILASTTPG 55  
 DB 1504 MTTPTTPASTTTLPTTTPSPPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1563  
 QY 56 ---VKSNLQPTTKTKNTT-----OTQPSKPTTKORONKPKPNNDHF 98  
 DB 1564 PPPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623  
 QY 99 FEVNFVPCISCNPTCWAICRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 158  
 DB 1624 TPI--TPPTSTTLPT-----TPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1675  
 QY 159 VPTTKPTERTNTTKNTNTTLLNTNTTGNPKLTSQMETHSTSEGNLSPSQVSTTSE 218  
 DB 1676 PPTTPSS--PITTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1731  
 QY 219 HPSQSPSPPTT 230  
 DB 1732 SPPTTTPPTT 1743

```

RESULT 14
MUC1_XENLA
ID MUC1_XENLA STANDARD; PRT; 662 AA.
AC Q05049;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integumentary mucin C.1 (PIM-C.1) (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6 AND 7).
RC TISSUE=Skin;
RX MEDLINE=93077556; PubMed=1447205;
RT Hauser F., Hoffmann W.;
RT "P-domains as shuffled cysteine-rich modules in integumentary mucin
RT C.1 (PIM-C.1) from Xenopus laevis. Polydispersity and genetic
RT polymorphism."
RL J. Biol. Chem. 267:24620-24624(1992).
CC -1- FUNCTION: Could be involved in defense against microbial
CC infections. Protects the epithelia from external environment.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=7;
CC Comment=Additional isoforms seem to exist. Experimental
CC confirmation may be lacking for some isoforms;
CC Name=1;
CC IsoId=Q05049-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q05049-2; Sequence=VSP_004650;
CC Name=3;
CC IsoId=Q05049-3; Sequence=VSP_004651;
CC Name=4;
CC IsoId=Q05049-4; Sequence=VSP_004647, VSP_004648;
CC Name=5;
CC IsoId=Q05049-5; Sequence=VSP_004646, VSP_004649, VSP_004650;
CC Name=6;
CC IsoId=Q05049-6; Sequence=VSP_004646, VSP_004648;
CC Name=7;
CC IsoId=Q05049-7; Sequence=VSP_004647;
CC -1- TISSUE SPECIFICITY: SKIN
CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED.
CC -1- SIMILARITY: Contains 6 P-type (trefoil) domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L02115; AAA74725.1; --
CC PIR: A45155; A45155.
CC HSSP: P01359; 2PSP.
CC InterPro: IPR000519; P_trefoil.
CC Pfam: PF00088; trefoil; 6.
CC PRINTS: PR00680; PTFREFOIL.
CC SMART: SM00018; PD; 6.
CC PROSITE: PS00025; P_TREFOIL; 6.
KW Repeat; Glycoprotein; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 81 144 8 X 8 AA APPROXIMATE TANDEM REPEATS,
FT REPEAT 81 88 ALA/THR-RICH.
FT REPEAT 89 96 1-1.
FT REPEAT 97 104 1-2.
FT REPEAT 105 112 1-3.
FT REPEAT 113 120 1-4.
FT REPEAT 113 120 1-5.

```

```

FT REPEAT 121 128
FT REPEAT 129 136
FT REPEAT 137 144
FT DOMAIN 161 202 P-TYPE 1.
FT DOMAIN 218 301 8 X APPROXIMATE TANDEM REPEATS, THR-RICH.
FT REPEAT 218 224 2-1.
FT REPEAT 225 239 2-2.
FT REPEAT 240 249 2-3.
FT REPEAT 250 259 2-4.
FT REPEAT 260 275 2-5.
FT REPEAT 276 287 2-6.
FT REPEAT 288 294 2-7.
FT REPEAT 295 301 2-8.
FT DOMAIN 306 347 P-TYPE 2.
FT DOMAIN 353 394 P-TYPE 3.
FT DOMAIN 402 522 12 X APPROXIMATE TANDEM REPEATS, THR-
FT RICH.
FT REPEAT 402 411 3-1.
FT REPEAT 412 419 3-2.
FT REPEAT 420 431 3-3.
FT REPEAT 432 443 3-4.
FT REPEAT 444 453 3-5.
FT REPEAT 454 460 3-6.
FT REPEAT 461 472 3-7.
FT REPEAT 473 479 3-8.
FT REPEAT 480 491 3-9.
FT REPEAT 492 498 3-10.
FT REPEAT 499 515 3-11.
FT REPEAT 516 522 3-12.
FT DOMAIN 525 566 P-TYPE 4.
FT DOMAIN 572 613 P-TYPE 5.
FT DOMAIN 620 661 P-TYPE 6.
FT DISULFID 162 188 BY SIMILARITY.
FT DISULFID 172 187 BY SIMILARITY.
FT DISULFID 182 199 BY SIMILARITY.
FT DISULFID 307 333 BY SIMILARITY.
FT DISULFID 317 332 BY SIMILARITY.
FT DISULFID 327 344 BY SIMILARITY.
FT DISULFID 354 380 BY SIMILARITY.
FT DISULFID 364 379 BY SIMILARITY.
FT DISULFID 374 391 BY SIMILARITY.
FT DISULFID 526 552 BY SIMILARITY.
FT DISULFID 536 551 BY SIMILARITY.
FT DISULFID 546 563 BY SIMILARITY.
FT DISULFID 573 599 BY SIMILARITY.
FT DISULFID 583 598 BY SIMILARITY.
FT DISULFID 593 610 BY SIMILARITY.
FT DISULFID 621 647 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 641 658 BY SIMILARITY.
FT VARSPLIC 240 259 Missing (in isoform 5 and isoform 6).
FT VARSPLIC 250 259 /FTID=VSP_004646.
FT VARSPLIC 276 294 Missing (in isoform 4 and isoform 7).
FT VARSPLIC 276 294 /FTID=VSP_004647.
FT VARSPLIC 276 294 Missing (in isoform 4 and isoform 6).
FT VARSPLIC 276 294 /FTID=VSP_004648.
FT VARSPLIC 276 294 Missing (in isoform 5).
FT VARSPLIC 306 350 /FTID=VSP_004649.
FT VARSPLIC 420 498 Missing (in isoform 2 and isoform 5).
FT VARSPLIC 420 498 /FTID=VSP_004650.
FT VARSPLIC 420 498 Missing (in isoform 3).
FT VARSPLIC 420 498 /FTID=VSP_004651.
FT VARIANT 276 276 K -> E.
FT VARIANT 354 354 C -> R.
FT VARIANT 415 415 T -> A.
SQ SEQUENCE 662 AA; 67774 MW; F085277F1ED2FD40 CRC64;
Query Match 15.5%; Score 191; DB 1; Length 662;
Best Local Similarity 29.0%; Pred. No. 5.3e-06;
Matches 80; Conservative 19; Mismatches 111; Indels 66; Gaps 12;
QY 2 KVTLTATIIQDATSIKNTT-PTYLTQDPQLGISFNSLSEITSQTTTLASTTPGVKSNL 60

```







C:Accession: A94048; A93599; A04039  
R:Wertz, G.W.; Collins, P.L.; Huang, Y.; Gruber, C.; Levine, S.; Ball, L.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4075-4079, 1985  
A:Title: Nucleotide sequence of the G protein gene of human respiratory syncytial virus  
A:Reference number: A94048; MUID:85216636; PMID:3858865  
A:Accession: A94048  
A:Molecule type: mRNA  
A:Residues: 1-298 <WER>  
A:Cross-references: GB:M11486; GB:K01459; GB:K02719; GB:K03348; GB:K03349; GB:M11217; GB:7.1; PID:G333932  
A:Note: residues 207-298 are identical with residues 376-467 of the nucleocapsid protein  
A:Note: this protein may carry 40-80 separate O-linked carbohydrate chains distributed a  
R:Satake, M.; Colligan, J.E.; Elango, N.; Norrby, E.; Venkatesan, S.  
Nucleic Acids Res. 13, 7795-7812, 1985  
A:Title: Respiratory syncytial virus envelope glycoprotein (G) has a novel structure.  
A:Reference number: A93599; MUID:86067198; PMID:4069997  
A:Accession: A93599  
A:Molecule type: mRNA  
A:Residues: 1-298 <SAT>  
A:Cross-references: GB:X03149; NID:G60997; PIDN:CAA26928.1; PID:G60998  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:38-66/Domain: transmembrane #status predicted <TM>  
F:85,103,135,237,251/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 91.1%; Score 1120; DB 1; Length 298;  
Best Local Similarity 92.7%; Pred. No. 1.6e-68; Indels 0; Gaps 0;  
Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
QY 1 HKVLTLLTAIIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTTTILASTTPGVKSNL 60  
DB 67 HKVPTTLLTAIIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTTTILASTTPGVKSTL 126  
QY 61 QPTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 246  
QY 181 LITNNTTGNPKLTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 231  
DB 247 LLTSNTTGNPELTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 298  
RESULT 3  
Attachment protein - human respiratory syncytial virus (strain RSB6256)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: JQ1208  
R:Can, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: JQ1208; MUID:91374005; PMID:1895054  
A:Accession: JQ1208  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 85.6%; Score 1052; DB 2; Length 297;  
Best Local Similarity 87.4%; Pred. No. 6.2e-64; Indels 0; Gaps 0;  
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;  
QY 1 HKVLTLLTAIIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTTTILASTTPGVKSNL 60  
DB 67 HKVLTLLTAIIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTTTILASTTPGVKSTL 126  
QY 61 QPTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 246  
QY 181 LITNNTTGNPKLTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 231  
DB 247 LLTSNTTGNPELTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 298  
RESULT 4  
Attachment protein - human respiratory syncytial virus (strain RSB1734)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: JQ1205  
R:Can, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: JQ1204; MUID:91374005; PMID:1895054  
A:Accession: JQ1205  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 85.0%; Score 1045; DB 2; Length 297;  
Best Local Similarity 86.6%; Pred. No. 1.8e-63; Indels 0; Gaps 0;  
Matches 200; Conservative 10; Mismatches 21; Indels 0; Gaps 0;  
QY 1 HKVLTLLTAIIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTTTILASTTPGVKSNL 60  
DB 67 HKVLTLLTAIIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTTTILASTTPGVKSTL 126  
QY 61 QPTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 246  
QY 181 LITNNTTGNPKLTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 231  
DB 247 LLTSNTTGNPELTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 297  
RESULT 5  
Attachment protein - human respiratory syncytial virus  
C:Species: human respiratory syncytial virus  
C>Date: 11-Nov-1997 #sequence\_revision 11-Nov-1997 #text\_change 26-Feb-1998  
C:Accession: JCS680  
R:Geng, X.; Wang, Z.; Qian, X.; Zhu, R.; Deng, J.; Du, J.; Zhu, Z.  
Chinese J. Virol. 12, 317-322, 1996  
A:Title: Molecular analysis of G protein gene of a respiratory syncytial virus strain is  
A:Reference number: JCS680  
A:Accession: JCS680  
A:Molecule type: mRNA  
A:Residues: 1-298 <GEN>  
A:Experimental source: strain B79  
A:Note: the authors translated the codon TTT for residue 165 and 170 as Glu, TTC for res  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
F:1-38/Domain: intracellular #status predicted <INT>  
F:39-66/Domain: transmembrane #status predicted <TM>

QY 61 QPTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 246  
QY 181 LITNNTTGNPKLTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 230  
DB 247 LLTSNTTGNPELTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 296  
RESULT 4  
Attachment protein - human respiratory syncytial virus (strain RSB1734)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: JQ1205  
R:Can, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: JQ1204; MUID:91374005; PMID:1895054  
A:Accession: JQ1205  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 85.0%; Score 1045; DB 2; Length 297;  
Best Local Similarity 86.6%; Pred. No. 1.8e-63; Indels 0; Gaps 0;  
Matches 200; Conservative 10; Mismatches 21; Indels 0; Gaps 0;  
QY 1 HKVLTLLTAIIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTTTILASTTPGVKSNL 60  
DB 67 HKVLTLLTAIIQDATSQIKNTTPTLTQDPQLGIFSLSLSEITSTTTILASTTPGVKSTL 126  
QY 61 QPTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKKPTKTTKDLKPQTTKPEVPTTKTEPTTNTTKNITTT 246  
QY 181 LITNNTTGNPKLTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 231  
DB 247 LLTSNTTGNPELTSMQETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT 297  
RESULT 5  
Attachment protein - human respiratory syncytial virus  
C:Species: human respiratory syncytial virus  
C>Date: 11-Nov-1997 #sequence\_revision 11-Nov-1997 #text\_change 26-Feb-1998  
C:Accession: JCS680  
R:Geng, X.; Wang, Z.; Qian, X.; Zhu, R.; Deng, J.; Du, J.; Zhu, Z.  
Chinese J. Virol. 12, 317-322, 1996  
A:Title: Molecular analysis of G protein gene of a respiratory syncytial virus strain is  
A:Reference number: JCS680  
A:Accession: JCS680  
A:Molecule type: mRNA  
A:Residues: 1-298 <GEN>  
A:Experimental source: strain B79  
A:Note: the authors translated the codon TTT for residue 165 and 170 as Glu, TTC for res  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
F:1-38/Domain: intracellular #status predicted <INT>  
F:39-66/Domain: transmembrane #status predicted <TM>



RESULT 9  
 JQ1204  
 attachment protein - human respiratory syncytial virus (strain RSB642)  
 N/Alternate names: G protein  
 C:Species: human respiratory syncytial virus  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C:Accession: JQ1204  
 R:Gane, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
 A:Reference number: JQ1204; MUID:91374005; PMID:1895054  
 A:Accession: JQ1204  
 A:Molecule type: mRNA  
 A:Residues: 1-297 <CAN>  
 C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill  
 children and adults.  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C:Keywords: glycoprotein; transmembrane protein  
 F:85,103,135,144,237,273/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 82.1%; Score 1009; DB 2; Length 297;  
 Best Local Similarity 84.0%; Pred. No. 4.8e-61;  
 Matches 194; Conservative 11; Mismatches 26; Indels 0; Gaps 0;  
 QY 1 HKVTLTALIQDANSQIKNTTPTLTQDPQLGIFSNLSITTSQTTLASTTPGVKSNL 60  
 DB 67 HRVTSITTLIQDANSQIKNTTPTLTQDPQLGIFSNLSITTSQTTLASTTPGVKLT 126  
 QY 61 QPTTVTKNTTTTQTPSKPTTKQKQKPPKPNNDPHEFVFNFCVCSNNPTCWAIC 120  
 DB 127 QSTTVRKNTTTTQAPKSKTTKQKQKPPKPNNDPHEFVFNFCVCSNNPTCWAIC 186  
 QY 121 KRIPNKKPGKTTTKPTKPTTKTKKLPQTKPKKEVPTTKPTKEPTINTTKNITT 180  
 DB 187 KRIPNKKPGKTTTKPTKPTTKTKKLPQTKPKKEVPTTKPTKEPTINTTKNITT 246  
 QY 181 LLTNNNTGNPKLTQSMETFTSTSEGNLSPSQVSTTSEHPSQSPNTR 231  
 DB 247 PLTNTARNPELTQSMETFTSTSEGNLSPSQVSTTSEHPSQSPNTPR 297  
 RESULT 10  
 MGNZ60  
 major surface glycoprotein G - human respiratory syncytial virus (strain 8/60)  
 N/Alternate names: attachment glycoprotein G  
 C:Species: human respiratory syncytial virus  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C:Accession: A37077  
 R:Sullender, W.M.; Anderson, K.; Wertz, G.W.  
 Virology 178, 195-203, 1990  
 A:Title: The respiratory syncytial virus subgroup B attachment glycoprotein: analysis of  
 gous subgroup virus challenge.  
 A:Reference number: A37077; MUID:90357765; PMID:1697126  
 A:Accession: A37077  
 A:Molecule type: mRNA  
 A:Residues: 1-292 <SUL>  
 A:Cross-references: EMBL:M55633; NID:G333944; PIDN:AAA47413.1; PID:G333945  
 C:Genetics: G  
 A:Gene: G  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C:Keywords: glycoprotein; transmembrane protein  
 F:45-63/Domain: transmembrane #status predicted <TN>  
 F:81,86,100,230,290/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 41.1%; Score 505.5; DB 1; Length 292;  
 Best Local Similarity 47.6%; Pred. No. 3.5e-27;  
 Matches 108; Conservative 28; Mismatches 88; Indels 3; Gaps 3;  
 QY 1 HKVTLTALIQDANSQIKNTTPTLTQDPQLGIFSNLSITTSQTTLASTTPGVKSNL 60  
 DB 67 HKVTLTAVTQTIKNHTGKNIISTYLTQVPPRVNNSKQPTTSPITNSATISPNKTS 126

QY 61 QPTTVTKNTTTTQTPSKPTTKQKQKPPKPNNDPHEFVFNFCVCSNNPTCWAIC 120  
 DB 127 HHTTAQTKGRITTTQTNTKPSKSKNPPKPKDDYHFEVFNFCVCSICGNQLCKSIC 186  
 QY 121 KRIPNKKPGKTTTKPTKPTTKTKKLPQTKPKKEVPTTKPTKEPTINTTKNITT 179  
 DB 187 KTIKPNKPKKTTTPTKPTTKTKKLPQTKPKKEVPTTKPTKEPTINTTKNITT 246  
 QY 180 TLLTNNNTGNPKLTQSMETFTSTSEGNLSPSQVSTTSEHPSQSPSP 226  
 DB 247 SQSTVLDITTPKYTIQQQSLHSTTSSTPSQTPTASE-PS-TSPN 291  
 RESULT 11  
 MGNZ18  
 major surface glycoprotein G - human respiratory syncytial virus (strain 18537)  
 C:Species: human respiratory syncytial virus  
 C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999  
 C:Accession: B32703  
 R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987  
 A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and B:  
 A:Reference number: A32703; MUID:87289657; PMID:2441388  
 A:Accession: B32703  
 A:Molecule type: mRNA  
 A:Residues: 1-292 <JOH>  
 A:Cross-references: GB:M17213; NID:G333942; PIDN:AAA47412.1; PID:G333943  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C:Keywords: glycoprotein; transmembrane protein  
 F:41-63/Domain: transmembrane #status predicted <TN>  
 F:81,86,100/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 40.4%; Score 497; DB 1; Length 292;  
 Best Local Similarity 46.8%; Pred. No. 1.3e-26;  
 Matches 104; Conservative 28; Mismatches 88; Indels 2; Gaps 2;  
 QY 1 HKVTLTALIQDANSQIKNTTPTLTQDPQLGIFSNLSITTSQTTLASTTPGVKSNL 60  
 DB 67 HKVTLTAVTQTIKNHTGKNIISTYLTQVPPRVNNSKQPTTSPITNSATISPNKTS 126  
 QY 61 QPTTVTKNTTTTQTPSKPTTKQKQKPPKPNNDPHEFVFNFCVCSNNPTCWAIC 120  
 DB 127 HHTTAQTKGRITTTQTNTKPSKSKNPPKPKDDYHFEVFNFCVCSICGNQLCKSIC 186  
 QY 121 KRIPNKKPGKTTTKPTKPTTKTKKLPQTKPKKEVPTTKPTKEPTINTTKNITT 179  
 DB 187 KTIKPNKPKKTTTPTKPTTKTKKLPQTKPKKEVPTTKPTKEPTINTTKNITT 246  
 QY 180 TLLTNNNTGNPKLTQSMETFTSTSEGNLSPSQVSTTSEHPS 221  
 DB 247 SQSTVLDITTPKYTIQQQSLHSTTSSTPSQTPTASE-PS 287  
 RESULT 12  
 VHNZ  
 nucleocapsid protein (version 2) - human respiratory syncytial virus  
 C:Species: human respiratory syncytial virus  
 C:Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 16-Jul-1999  
 C:Accession: A04026  
 R:Elango, N.; Venkatesan, S.  
 Nucleic Acids Res. 11, 5941-5951, 1983  
 A:Title: Amino acid sequence of respiratory syncytial virus capsid protein.  
 A:Reference number: A04026; MUID:83299261; PMID:6310521  
 A:Accession: A04026  
 A:Molecule type: mRNA  
 A:Residues: 1-467 <ELA>  
 A:Cross-references: GB:X00001; NID:G61215; PIDN:CAA24906.1; PID:G61216  
 C:Genetics: N  
 A:Gene: N  
 C:Superfamily: respiratory syncytial virus nucleocapsid protein  
 C:Keywords: nucleocapsid  
 Query Match 34.6%; Score 425; DB 1; Length 467;



Search completed: October 29, 2003, 17:41:50  
Job time : 23.3245 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:41:00 ; Search time 40.7094 Seconds  
(without alignments)  
975.349 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229  
Sequence: 1 HKVLTITAIIDATSIQIKNT.....VSTTSEHPSPSPNTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 642050 seqs, 171146064 residues

Total number of hits satisfying chosen parameters: 642050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

```

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 193   | 15.7        | 5179   | 9  | US-09-922-217-1068 |
| 2          | 193   | 15.7        | 5179   | 10 | US-09-833-263-1068 |
| 3          | 193   | 15.7        | 5179   | 14 | US-10-025-380-1068 |
| 4          | 169.5 | 13.8        | 1367   | 10 | US-09-801-368-108  |
| 5          | 162   | 13.2        | 941    | 14 | US-10-124-557-14   |
| 6          | 162   | 13.2        | 1022   | 14 | US-10-124-557-84   |
| 7          | 162   | 13.2        | 1038   | 14 | US-10-124-557-74   |
| 8          | 162   | 13.2        | 1049   | 14 | US-10-124-557-58   |
| 9          | 162   | 13.2        | 1140   | 14 | US-10-124-557-104  |
| 10         | 162   | 13.2        | 1270   | 14 | US-10-124-557-44   |
| 11         | 162   | 13.2        | 1311   | 14 | US-10-124-557-42   |
| 12         | 162   | 13.2        | 1313   | 14 | US-10-124-557-142  |
| 13         | 162   | 13.2        | 1314   | 14 | US-10-124-557-50   |
| 14         | 162   | 13.2        | 1320   | 14 | US-10-124-557-46   |
| 15         | 162   | 13.2        | 1320   | 14 | US-10-124-557-60   |

|    |       |      |      |    |                     |
|----|-------|------|------|----|---------------------|
| 16 | 162   | 13.2 | 1354 | 14 | US-10-124-557-48    |
| 17 | 162   | 13.2 | 1361 | 14 | US-10-124-557-40    |
| 18 | 162   | 13.2 | 1363 | 14 | US-10-124-557-52    |
| 19 | 162   | 13.2 | 1404 | 14 | US-10-124-557-2     |
| 20 | 162   | 13.2 | 1404 | 14 | US-10-124-557-62    |
| 21 | 161.5 | 13.1 | 288  | 9  | US-09-216-393-341   |
| 22 | 161.5 | 13.1 | 288  | 9  | US-09-216-393-344   |
| 23 | 161.5 | 13.1 | 288  | 12 | US-10-321-856-341   |
| 24 | 161.5 | 13.1 | 288  | 12 | US-10-321-856-344   |
| 25 | 161.5 | 13.1 | 800  | 12 | US-10-029-386-32198 |
| 26 | 152.5 | 12.4 | 538  | 12 | US-10-038-694-3     |
| 27 | 146   | 11.9 | 528  | 12 | US-09-840-746-20    |
| 28 | 144.5 | 11.8 | 1260 | 15 | US-10-245-802-8     |
| 29 | 144.5 | 11.8 | 2828 | 10 | US-09-985-129-21    |
| 30 | 144.5 | 11.8 | 2828 | 10 | US-09-991-630-21    |
| 31 | 144.5 | 11.8 | 2828 | 12 | US-10-301-822-49    |
| 32 | 144.5 | 11.8 | 2828 | 12 | US-10-032-189-126   |
| 33 | 144.5 | 11.8 | 2828 | 15 | US-10-176-847-54    |
| 34 | 144.5 | 11.8 | 2828 | 15 | US-10-177-293-110   |
| 35 | 143   | 11.6 | 995  | 11 | US-09-984-130-48    |
| 36 | 143   | 11.6 | 995  | 12 | US-09-836-353A-48   |
| 37 | 142.5 | 11.6 | 449  | 12 | US-10-137-870-224   |
| 38 | 142.5 | 11.6 | 449  | 12 | US-10-140-018-224   |
| 39 | 142.5 | 11.6 | 449  | 12 | US-10-140-021-224   |
| 40 | 142.5 | 11.6 | 449  | 12 | US-10-140-274-224   |
| 41 | 142.5 | 11.6 | 449  | 12 | US-10-140-471-224   |
| 42 | 142.5 | 11.6 | 449  | 12 | US-10-140-807-224   |
| 43 | 142.5 | 11.6 | 449  | 12 | US-10-140-922-224   |
| 44 | 142.5 | 11.6 | 449  | 12 | US-10-140-924-224   |
| 45 | 142.5 | 11.6 | 449  | 12 | US-10-140-926-224   |

#### ALIGNMENTS

#### RESULT 1

```

US-09-922-217-1068
; Sequence 1068, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSTICS
; FILE REFERENCE: 210121.471C13
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1068

```

Query Match 15.7%; Score 193; DB 9; Length 5179;  
Best Local Similarity 30.8%; Pred. No. 1.6e-06;  
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;  
Cv 5 LTTAIQDA--TSQIKNTTP-----TYLTQDPOLGIFSNNLSKIT--SOTTILASTPG 55  
Db 1504 MTTPIFPASTTTLPLPTTTPSPPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1563

```
QY 56 ----VKSNLQPTTVKTKNTTT-----QTQPSKPTTKQKQKPNKPNNDHF 98
Db 1564 PPPTTTTTPPTTSPPTTSPPTTITTTTTPPTTSPPTTITTTTTPPTTSPPTT 1623
QY 99 FEVNFVPCISNNPTCWAICRIPNKKPGKTKTKTKKPKTKTKKDLKPQTKPKKE 158
Db 1624 TPI--TPPTSTTLPLPT-----TTPSPPTTITTTTTPPTTSPPTTITTTTTP 1675
QY 159 VPTTKTEPTINTTKNTITLLTNNTGKPKLTQSMETHSTSSGNLSPSQVSTTSE 218
Db 1676 PPTTTPSS--PITTPSPPTTITMTPTSP--SSPITTTTTPSSTTTTTPSPPTTMTTP 1731
QY 219 HPSQSPSPNTT 230
Db 1732 SPPTTSPPTT 1743

RESULT 2
US-09-833-263-1068
; Sequence 1068, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jonathan D.
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Meagher, Madeleine J.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1068

Query Match 15.7%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 1.6e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 5 LTTAIIQDA--TSQIKNTTP-----TYLTQDPQLGISFNSLSEIT--SQTTILASTTPG 55
Db 1504 MTPTTPPASTTTLPLPTTTPSPPTTITTTTTPPTTTPPTTTPPTTTPPTTTP 1563
QY 56 ----VKSNLQPTTVKTKNTTT-----QTQPSKPTTKQKQKPNKPNNDHF 98
Db 1564 PPPTTTTTPPTTSPPTTSPPTTITTTTTPPTTSPPTTITTTTTPPTTSPPTT 1623
QY 99 FEVNFVPCISNNPTCWAICRIPNKKPGKTKTKTKKPKTKTKKDLKPQTKPKKE 158
Db 1624 TPI--TPPTSTTLPLPT-----TTPSPPTTITTTTTPPTTTPPTTTPPTTTP 1675
QY 159 VPTTKTEPTINTTKNTITLLTNNTGKPKLTQSMETHSTSSGNLSPSQVSTTSE 218
Db 1676 PPTTTPSS--PITTPSPPTTITMTPTSP--SSPITTTTTPSSTTTTTPSPPTTMTTP 1731
QY 219 HPSQSPSPNTT 230
Db 1732 SPPTTSPPTT 1743

RESULT 3
US-10-025-380-1068
; Sequence 1068, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
```

```
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1068

Query Match 15.7%; Score 193; DB 14; Length 5179;
Best Local Similarity 30.6%; Pred. No. 1.6e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 5 LTTAIIQDA--TSQIKNTTP-----TYLTQDPQLGISFNSLSEIT--SQTTILASTTPG 55
Db 1504 MTPTTPPASTTTLPLPTTTPSPPTTITTTTTPPTTTPPTTTPPTTTPPTTTP 1563
QY 56 ----VKSNLQPTTVKTKNTTT-----QTQPSKPTTKQKQKPNKPNNDHF 98
Db 1564 PPPTTTTTPPTTSPPTTSPPTTITTTTTPPTTTPPTTITTTTTPPTTTPPTT 1623
QY 99 FEVNFVPCISNNPTCWAICRIPNKKPGKTKTKTKKPKTKTKKDLKPQTKPKKE 158
Db 1624 TPI--TPPTSTTLPLPT-----TTPSPPTTITTTTTPPTTTPPTTTPPTTTP 1675
QY 159 VPTTKTEPTINTTKNTITLLTNNTGKPKLTQSMETHSTSSGNLSPSQVSTTSE 218
Db 1676 PPTTTPSS--PITTPSPPTTITMTPTSP--SSPITTTTTPSSTTTTTPSPPTTMTTP 1731
QY 219 HPSQSPSPNTT 230
Db 1732 SPPTTSPPTT 1743

RESULT 4
US-09-801-368-108
; Sequence 108, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250Alman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
```





```
;
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1022 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-10-124-557-84

Query Match      13.2%; Score 162; DB 14; Length 1022;
Best Local Similarity 24.6%; Pred. No. 7.8e-05;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY      6 TTAIIQDATSQIKNTTPTVLTQDPQLGISFNSLSEITSQTITLSTTPGVKSNLQPTTV 65
Db      206 TSKVLAKPTPKAETTK-----GPALTTPKEPTPTTPKEPASTTP---KEPTPTTI 253
QY      66 K-----TKNTTTTQTPSKPTTKQKQKPPKPNNDHFHFEVNFVPCISCSN 112
Db      254 KSAPPTPKAPATTTKSAPTTKEPATTTPKEPATTTPKEPATTTPKEPATTTPKEP 302
QY      113 NPTCWAICKRIPN-----KKPGKKTTPK-----TKKPTFKTK 146
Db      303 APPTTKSAPTTKEPATTTPKDPATTPKEPATTTPKEPATTTPKEPATTTPKEPATTTPK 362
QY      147 KDL-----KPQTTKPEVPTTKP-----TEPTINTKNTITLLTNNITGNPKLTS 194
Db      363 EPAPTAPKPAPTTPKEPATTTPKEPATTTPKEPATTTPKEPATTTPKSAPTTKEPAPT 422
QY      195 QMETFHSTSEGNLSFQSVSTTSEHPSQSPSPNTT 230
Db      423 TTKSAPTTKE-----PSPTTKKEPATTTPKEPATT 454

RESULT 7
US-10-124-557-74
; Sequence 74, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
```

```
;
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1038 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-124-557-74

Query Match      13.2%; Score 162; DB 14; Length 1038;
Best Local Similarity 24.6%; Pred. No. 7.9e-05;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY      6 TTAIIQDATSQIKNTTPTVLTQDPQLGISFNSLSEITSQTITLSTTPGVKSNLQPTTV 65
Db      190 TSKVLAKPTPKAETTK-----GPALTTPKEPTPTTPKEPASTTP---KEPTPTTI 237
QY      66 K-----TKNTTTTQTPSKPTTKQKQKPPKPNNDHFHFEVNFVPCISCSN 112
Db      238 KSAPTTKEPATTTPKSAPTTKEPATTTPKEPATTTPKEPATTTPKEPATTTPKEP 286
QY      113 NPTCWAICKRIPN-----KKPGKKTTPK-----TKKPTFKTK 146
Db      287 APPTTKSAPTTKEPATTTPKDPATTPKEPATTTPKEPATTTPKEPATTTPKEPATTTPK 346
QY      147 KDL-----KPQTTKPEVPTTKP-----TEPTINTKNTITLLTNNITGNPKLTS 194
Db      347 EPAPTAPKPAPTTPKEPATTTPKEPATTTPKEPATTTPKEPATTTPKSAPTTKEPAPT 406
QY      195 QMETFHSTSEGNLSFQSVSTTSEHPSQSPSPNTT 230
Db      407 TTKSAPTTKE-----PSPTTKKEPATTTPKEPATT 438

RESULT 8
US-10-124-557-58
; Sequence 58, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
```

FILED DATE: 16-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
APPLICATION NUMBER: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 876-1170  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1049 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 58:

US-10-124-557-58  
Query Match 13.2%; Score 162; DB 14; Length 1049;  
Best Local Similarity 24.6%; Pred. No. 8e-05;  
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;  
QY 6 TTAIQDQTSQIKNTPTVLTQDPQLGISFNSLSITTSQTTILASTTPGVKSNLQPTTV 65  
Db 233 TSKVLAKTPKAEITK-----GPAITPKPEPTTPKPEASTTP---KEPTPTTI 280  
QY 66 K-----TKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFVPCISCSN 112  
Db 281 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKPEAPTTKPEAPTTKPEAPTTKPE 329  
QY 113 NPTCWAICRIPN-----KPGKKTTPK-----TKKPTFTTK 146  
Db 330 APTTKSAPTTPKEPAPTTKPAPTTPKEPAPTTKPEPTTPKPEAPTTKPEAPTTK 389  
QY 147 KDL-----KPQTKPEVPTTKP-----TEEPINTTKNTITLLTNNTGNPKLTS 194  
Db 330 EPAPAPKKEPAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTKPE 449  
QY 195 QMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTT 230  
Db 450 TTKSAPTTPKE-----PSPTTKPEAPTTKPEAPTT 481

RESULT 9  
US-10-124-557-104  
Sequence 104, Application US/10124557  
Publication No. US20020137894A1  
GENERAL INFORMATION:  
APPLICANT: Turner, Katherine  
Clark, Stephen C.  
Jacobs, Kenneth  
Hewick, Rodney M.  
Gesner, Thomas G.

TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/124,557  
FILING DATE: 16-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
APPLICATION NUMBER: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserr, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 876-1170  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 104:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1140 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 104:  
US-10-124-557-104

Query Match 13.2%; Score 162; DB 14; Length 1140;  
Best Local Similarity 24.6%; Pred. No. 8.9e-05;  
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;  
QY 6 TTAIQDQTSQIKNTPTVLTQDPQLGISFNSLSITTSQTTILASTTPGVKSNLQPTTV 65  
Db 324 TSKVLAKTPKAEITK-----GPAITPKPEPTTPKPEASTTP---KEPTPTTI 371  
QY 66 K-----TKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFVPCISCSN 112  
Db 372 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKPEAPTTKPEAPTTKPEAPTTKPE 420  
QY 113 NPTCWAICRIPN-----KPGKKTTPK-----TKKPTFTTK 146  
Db 421 APTTKSAPTTPKEPAPTTKPAPTTPKEPAPTTKPEPTTPKPEAPTTKPEAPTTK 480  
QY 147 KDL-----KPQTKPEVPTTKP-----TEEPINTTKNTITLLTNNTGNPKLTS 194  
Db 481 EPAPAPKKEPAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTKPE 540  
QY 195 QMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTT 230  
Db 541 TTKSAPTTPKE-----PSPTTKPEAPTTKPEAPTT 572

RESULT 10  
US-10-124-557-44  
Sequence 44, Application US/10124557  
Publication No. US20020137894A1  
GENERAL INFORMATION:  
APPLICANT: Turner, Katherine  
Clark, Stephen C.  
Jacobs, Kenneth  
Hewick, Rodney M.  
Gesner, Thomas G.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge

STATE: Massachusetts  
 COUNTRY: U.S.A.  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/124,557  
 FILING DATE: 16-Apr-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/643,502  
 FILING DATE: 18-JAN-1991  
 APPLICATION NUMBER: US 07/546,114  
 FILING DATE: 29-JUN-1990  
 APPLICATION NUMBER: US 07/457,196  
 FILING DATE: 29-DEC-1989  
 APPLICATION NUMBER: US 07/390,901  
 FILING DATE: 08-AUG-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cseri, Luann  
 REGISTRATION NUMBER: 31,822  
 REFERENCE/DOCKET NUMBER: GI 5190  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)876-1170  
 TELEFAX: (617)876-5851  
 INFORMATION FOR SEQ ID NO: 44:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1270 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 44:  
 US-10-124-557-44

Query Match 13.28; Score 162; DB 14; Length 1270;  
 Best Local Similarity 24.68; Pred. No. 0.0001;  
 Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;  
 QY 6 TTAIIQDATSQIKNTTPTTLTQDQGLGSPSNLSEITSTQTTILASTTPGVKSNLQPTTV 65  
 DB 190 TSKVLAKPTPKAETTK-----GPAULTTKEPTTTPKEPATTP---KEPTPTTI 237  
 QY 66 K-----TKNTTTQTPSKPTTKQKQKPNKPNNDFFHEVFNFPVCSN 112  
 DB 238 KSAPTTPKEPATTTKSAPTTPKEPATTTTKEPATTPKEP-----APTTPKEP 286  
 QY 113 NPTCWAICKRIEN-----KKPKKTTTKP-----TKKPTTKTK 146  
 DB 287 APTTKSAPTTPKEPATTTKKAAPTTPKEPATTPKEPTTTPKEPATTPKEPATTPK 346  
 QY 147 KDL-----KPQTTKKEVPTTKP-----TEPTINTTKNITNTLLTNTTGNPKLTS 194  
 DB 347 EPAPTAPKKAAPTTPKEPATTPKEPATTTKPSPTTPKEPATTTKSAPTTPKEPAT 406  
 QY 195 QMETFHSTSSGNLSPSQVSTTSEHPSQSPSPNTT 230  
 DB 407 TTKSAPTTPKE-----PSPTTKKEPATTPKEPATTT 438

RESULT 11  
 US-10-124-557-42  
 ; Sequence 42, Application US/10124557  
 ; Publication No. US20020137894A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Turner, Katherine  
 ; Clark, Stephen C.  
 ; Jacobs, Kenneth  
 ; Gesner, Rodney M.  
 ; Gesner, Thomas G.  
 ; TITLE OF INVENTION: Megakaryocyte Stimulating Factors

NUMBER OF SEQUENCES: 143  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genetics Institute, Inc.  
 STREET: 87 CambridgePark Drive  
 CITY: Cambridge  
 STATE: Massachusetts  
 COUNTRY: U.S.A.  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/124,557  
 FILING DATE: 16-Apr-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/643,502  
 FILING DATE: 18-JAN-1991  
 APPLICATION NUMBER: US 07/546,114  
 FILING DATE: 29-JUN-1990  
 APPLICATION NUMBER: US 07/457,196  
 FILING DATE: 29-DEC-1989  
 APPLICATION NUMBER: US 07/390,901  
 FILING DATE: 08-AUG-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cseri, Luann  
 REGISTRATION NUMBER: 31,822  
 REFERENCE/DOCKET NUMBER: GI 5190  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)876-1170  
 TELEFAX: (617)876-5851  
 INFORMATION FOR SEQ ID NO: 42:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1311 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 42:  
 US-10-124-557-42

Query Match 13.24; Score 162; DB 14; Length 1311;  
 Best Local Similarity 24.68; Pred. No. 0.0001;  
 Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;  
 QY 6 TTAIIQDATSQIKNTTPTTLTQDQGLGSPSNLSEITSTQTTILASTTPGVKSNLQPTTV 65  
 DB 231 TSKVLAKPTPKAETTK-----GPAULTTKEPTTTPKEPATTP---KEPTPTTI 278  
 QY 66 K-----TKNTTTQTPSKPTTKQKQKPNKPNNDFFHEVFNFPVCSN 112  
 DB 279 KSAPTTPKEPATTTKSAPTTPKEPATTTTKEPATTPKEP-----APTTPKEP 327  
 QY 113 NPTCWAICKRIEN-----KKPKKTTTKP-----TKKPTTKTK 146  
 DB 328 APTTKSAPTTPKEPATTTKKAAPTTPKEPATTPKEPTTTPKEPATTPKEPATTPK 387  
 QY 147 KDL-----KPQTTKKEVPTTKP-----TEPTINTTKNITNTLLTNTTGNPKLTS 194  
 DB 388 EPAPTAPKKAAPTTPKEPATTPKEPATTTKPSPTTPKEPATTTKSAPTTPKEPAT 447  
 QY 195 QMETFHSTSSGNLSPSQVSTTSEHPSQSPSPNTT 230  
 DB 448 TTKSAPTTPKE-----PSPTTKKEPATTPKEPATTT 479

RESULT 12  
 US-10-124-557-142  
 ; Sequence 142, Application US/10124557  
 ; Publication No. US20020137894A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Turner, Katherine

Clark, Stephen C.  
Jacobs, Kenneth  
Hewick, Rodney M.  
Gesner, Thomas G.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/124,557  
FILING DATE: 16-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
APPLICATION NUMBER: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserri, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 142:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 142:

Query Match 13.2%; Score 162; DB 14; Length 1313;  
Best Local Similarity 24.6%; Pred. No. 0.00011;  
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

|    |     |   |     |
|----|-----|---|-----|
| QY | 6   | TTAIIQDATSQIKNTTPTVLTQDPQLGISFSLSEITSTTTTILASTTPGVKSNLQPTTV | 65  |
| Db | 233 | TSKVLAKTPKAEITTK-----GPAITTKPEPTTTPKEPASTTP---KEPTPTTI      | 280 |
| QY | 66  | K-----TKNTTTTQPSKPTTKQKPNKPNNDFFHVFNFVPCISCN                | 112 |
| Db | 281 | KSAPTTKPEPAPTTKSAPTTKPEPAPTTTKPEPAPTTKPEP-----APTTKEP       | 329 |
| QY | 113 | NPTCWAICKRIPN-----KPGKKTTPK-----TKKPTFKTK                   | 146 |
| Db | 330 | APTTKSAPTTKPEPAPTTKPEPAPTTKPEPAPTTKPEPAPTTKPEPAPTTK         | 389 |
| QY | 147 | KDL-----KPQTTKPEVPTTKP-----TEPTINTTKNITTTLLTNNTGNPKLTS      | 194 |
| Db | 390 | EPAPTAPKPAPTTKPEPAPTTKPEPAPTTTKPEPAPTTKPEPAPTTKSAPTTK       | 449 |
| QY | 195 | QMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT                          | 230 |
| Db | 450 | TTKSAPTTKPE-----PSPTTKPEPAPTTKPEPAPTT                       | 481 |

RESULT 13

US-10-124-557-50  
Sequence 50, Application US/10124557  
Publication No. US20020137894A1  
GENERAL INFORMATION:  
APPLICANT: Turner, Katherine  
Clark, Stephen C.  
Jacobs, Kenneth  
Hewick, Rodney M.  
Gesner, Thomas G.  
TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
NUMBER OF SEQUENCES: 143  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/124,557  
FILING DATE: 16-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,502  
FILING DATE: 18-JAN-1991  
APPLICATION NUMBER: US 07/546,114  
FILING DATE: 29-JUN-1990  
APPLICATION NUMBER: US 07/457,196  
FILING DATE: 29-DEC-1989  
APPLICATION NUMBER: US 07/390,901  
FILING DATE: 08-AUG-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Cserri, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: GI 5190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)876-1170  
TELEFAX: (617)876-5851  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1314 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:

Query Match 13.2%; Score 162; DB 14; Length 1314;  
Best Local Similarity 24.6%; Pred. No. 0.00011;  
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

|    |     |   |     |
|----|-----|---|-----|
| QY | 6   | TTAIIQDATSQIKNTTPTVLTQDPQLGISFSLSEITSTTTTILASTTPGVKSNLQPTTV | 65  |
| Db | 234 | TSKVLAKTPKAEITTK-----GPAITTKPEPTTTPKEPASTTP---KEPTPTTI      | 281 |
| QY | 66  | K-----TKNTTTTQPSKPTTKQKPNKPNNDFFHVFNFVPCISCN                | 112 |
| Db | 282 | KSAPTTKPEPAPTTKSAPTTKPEPAPTTTKPEPAPTTKPEP-----APTTKEP       | 330 |
| QY | 113 | NPTCWAICKRIPN-----KPGKKTTPK-----TKKPTFKTK                   | 146 |
| Db | 331 | APTTKSAPTTKPEPAPTTKPEPAPTTKPEPAPTTKPEPAPTTKPEPAPTTK         | 390 |
| QY | 147 | KDL-----KPQTTKPEVPTTKP-----TEPTINTTKNITTTLLTNNTGNPKLTS      | 194 |
| Db | 391 | EPAPTAPKPAPTTKPEPAPTTKPEPAPTTTKPEPAPTTKPEPAPTTKSAPTTK       | 450 |
| QY | 195 | QMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTT                          | 230 |

[illegible]

```
Qy 113 NPTCWAICKRIPN-----KKPGKXITTKP-----TKKPTFKTK 146
Db 337 APTTTKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPK 396
Qy 147 KDL-----KQOTYKPEVPTTKP-----TEPTINTKTNTITLLTNNITGNPKLTS 194
Db 397 EPAPTAPKKPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPT 456
Qy 195 QMETFHSSTSGNLSPOVSTTSEHPSOPSPPNNTT 230
Db 457 TTKSAPTTPKE-----PSFTTIKEPAPTTPKEPAPT 488
```

Search completed: October 29, 2003, 17:54:27  
Job time : 41.7094 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:36:05 ; Search time 12.6943 Seconds  
(without alignments)  
773.267 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVTLTAIRQDATSIKNT.....VSTTSHPSQSPSPNTTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgm2\_6/ptodata/2/iaa/5A COMB pep.\*  
2: /cgm2\_6/ptodata/2/iaa/5B COMB pep.\*  
3: /cgm2\_6/ptodata/2/iaa/6A COMB pep.\*  
4: /cgm2\_6/ptodata/2/iaa/6B COMB pep.\*  
5: /cgm2\_6/ptodata/2/iaa/PCTUS COMB pep.\*  
6: /cgm2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 1229  | 100.0       | 298    | 2     | US-08-467-963C-8  |
| 2          | 1229  | 100.0       | 298    | 2     | US-08-838-189D-8  |
| 3          | 1229  | 100.0       | 298    | 3     | US-08-852-344D-8  |
| 4          | 1229  | 100.0       | 298    | 3     | US-08-344-639E-8  |
| 5          | 1229  | 100.0       | 298    | 3     | US-08-467-961A-8  |
| 6          | 1229  | 100.0       | 298    | 3     | US-08-467-961A-8  |
| 7          | 1229  | 100.0       | 298    | 3     | US-08-001-554A-8  |
| 8          | 901   | 73.3        | 681    | 6     | 5194595-19        |
| 9          | 222   | 18.1        | 37     | 3     | US-08-793-792-12  |
| 10         | 193   | 15.7        | 32     | 3     | US-08-793-792-8   |
| 11         | 188   | 15.3        | 216    | 3     | US-08-928-361B-8  |
| 12         | 188   | 15.3        | 216    | 4     | US-08-928-361B-8  |
| 13         | 188   | 15.3        | 1837   | 4     | US-08-928-361B-5  |
| 14         | 188   | 15.3        | 1837   | 4     | US-08-588-995A-5  |
| 15         | 187.5 | 15.3        | 1721   | 3     | US-08-700-651-5   |
| 16         | 187.5 | 15.3        | 1721   | 3     | US-08-928-361B-6  |
| 17         | 187.5 | 15.3        | 1721   | 4     | US-08-588-995A-6  |
| 18         | 181.5 | 14.8        | 216    | 3     | US-08-928-361B-27 |
| 19         | 174   | 14.2        | 28     | 3     | US-08-793-792-4   |
| 20         | 173.5 | 14.1        | 249    | 3     | US-08-700-651-15  |
| 21         | 173.5 | 14.1        | 249    | 3     | US-08-928-361B-20 |
| 22         | 173.5 | 14.1        | 249    | 4     | US-08-588-995A-20 |
| 23         | 170.5 | 13.9        | 263    | 5     | PCT-US91-08177-13 |
| 24         | 164.5 | 13.4        | 2476   | 2     | US-08-276-967-2   |
| 25         | 162   | 13.2        | 941    | 4     | US-07-757-022B-14 |
| 26         | 162   | 13.2        | 1022   | 4     | US-07-757-022B-84 |
| 27         | 162   | 13.2        | 1038   | 4     | US-07-757-022B-74 |

## ALIGNMENTS

### RESULT 1

US-08-467-963C-8

; Sequence 8, Application US/08467963C

; Patent No. 5968776

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michael H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHYN, Mary E

; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES PROTECTION AGAINST

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES PROTECTION AGAINST

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6th Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/467,963C

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/838,189

; FILING DATE: 16-APR-1997

; APPLICATION NUMBER: US 08/001,554

; FILING DATE: 06-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9200117.1

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 298 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-467-963C-8

Sequence 58, Appl  
Sequence 104, Appl  
Sequence 44, Appl  
Sequence 42, Appl  
Sequence 142, Appl  
Sequence 50, Appl  
Sequence 46, Appl  
Sequence 60, Appl  
Sequence 48, Appl  
Sequence 40, Appl  
Sequence 52, Appl  
Sequence 2, Appl  
Sequence 62, Appl  
Sequence 341, Appl  
Sequence 344, Appl  
Sequence 12, Appl  
Sequence 17, Appl  
Sequence 17, Appl

Query Match 100.0%; Score 1229; DB 2; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 3e-105;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HKVLTLLAIQDATSQIKNTTPTLYTQDPQLGISFNSLSEITSTTTTILASTTTPGVKSNL 60  
 Db 67 HKVLTLLAIQDATSQIKNTTPTLYTQDPQLGISFNSLSEITSTTTTILASTTTPGVKSNL 126

Qy 61 OPTTVKTKNTTTTQTOPSKPTTKQKQKPNKPNNDHFEVFNVPVPCISCSNNPTCWAIC 120  
 Db 127 OPTTVKTKNTTTTQTOPSKPTTKQKQKPNKPNNDHFEVFNVPVPCISCSNNPTCWAIC 186

Qy 121 KRIPNKPKGKTKTKPTKPKTKTKDLKPKQTTKPKKEVPTTKTEPTTNTTKNTTTT 180  
 Db 187 KRIPNKPKGKTKTKPKTKTKDLKPKQTTKPKKEVPTTKTEPTTNTTKNTTTT 246

Qy 181 LLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 Db 247 LLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298

RESULT 2  
 US-08-838-189D-8  
 ; Sequence 8, Application US/08838189D  
 ; Patent No. 5998169  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KLEIN, Michel H  
 ; APPLICANT: DU, Run-Pan  
 ; APPLICANT: EWASYSHYN, Mary E  
 ; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
 ; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST  
 ; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
 ; NUMBER OF SEQUENCES: 38  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sim & McBurney  
 ; STREET: 6th Floor, 330 University Avenue  
 ; CITY: Toronto  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: M5G 1R7  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/838,189D  
 ; FILING DATE: 16-APR-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/001,554  
 ; FILING DATE: 06-JAN-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9200117.1  
 ; FILING DATE: 06-JAN-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: STEWART, Michael I  
 ; REGISTRATION NUMBER: 24,973  
 ; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (416) 595-1155  
 ; TELEFAX: (416) 595-1163  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 298 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-838-189D-8

Query Match 100.0%; Score 1229; DB 2; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 3e-105;

Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HKVLTLLAIQDATSQIKNTTPTLYTQDPQLGISFNSLSEITSTTTTILASTTTPGVKSNL 60  
 Db 67 HKVLTLLAIQDATSQIKNTTPTLYTQDPQLGISFNSLSEITSTTTTILASTTTPGVKSNL 126

Qy 61 OPTTVKTKNTTTTQTOPSKPTTKQKQKPNKPNNDHFEVFNVPVPCISCSNNPTCWAIC 120  
 Db 127 OPTTVKTKNTTTTQTOPSKPTTKQKQKPNKPNNDHFEVFNVPVPCISCSNNPTCWAIC 186

Qy 121 KRIPNKPKGKTKTKPTKPKTKTKDLKPKQTTKPKKEVPTTKTEPTTNTTKNTTTT 180  
 Db 187 KRIPNKPKGKTKTKPKTKTKDLKPKQTTKPKKEVPTTKTEPTTNTTKNTTTT 246

Qy 181 LLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 Db 247 LLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298

RESULT 3  
 US-08-852-344D-8  
 ; Sequence 8, Application US/08852344D  
 ; Patent No. 6017539  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KLEIN, Michel H  
 ; APPLICANT: DU, Run-Pan  
 ; APPLICANT: EWASYSHYN, Mary E  
 ; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION  
 ; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY  
 ; TITLE OF INVENTION: SYNCYTIAL VIRUS  
 ; NUMBER OF SEQUENCES: 37  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sim & McBurney  
 ; STREET: 6th Floor, 330 University Avenue  
 ; CITY: Toronto  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: M5G 1R7  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/852,344D  
 ; FILING DATE: 07-MAY-1997  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/344,639  
 ; FILING DATE: 14-NOV-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9200117.1  
 ; FILING DATE: 06-JAN-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: STEWART, Michael I  
 ; REGISTRATION NUMBER: 24,973  
 ; REFERENCE/DOCKET NUMBER: 1038-688 MIS:jb  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (416) 595-1155  
 ; TELEFAX: (416) 595-1163  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 298 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-852-344D-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 3e-105;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HKVLTLLAIQDATSQIKNTTPTLYTQDPQLGISFNSLSEITSTTTTILASTTTPGVKSNL 60

Db 67 HKVLTATAIQDATSQIKNTTPTVLTQDPQLGISFNSLSEITSTQTTILASTTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTPSKPTTKQRONKPNKPNNDFFHFEVFNVPSCISNNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTPSKPTTKQRONKPNKPNNDFFHFEVFNVPSCISNNPTCWAIC 186  
QY 121 KRIPNKPFGKTKTKPTKPTFKTKKDLKPQTKPKKEVPTTKTEBPTINTTKNTITTT 180  
Db 187 KRIPNKPFGKTKTKPTKPTFKTKKDLKPQTKPKKEVPTTKTEBPTINTTKNTITTT 246  
QY 181 LLTNNTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
Db 247 LLTNNTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298

## RESULT 4

US-08-344-639E-8  
; Sequence 8, Application US/08344639E  
; Patent No. 6033668  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasyszyn, Mary E  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS  
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
; TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/344,639E  
; FILING DATE: 14-NOV-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-391 MTS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-344-639E-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTATAIQDATSQIKNTTPTVLTQDPQLGISFNSLSEITSTQTTILASTTTPGVKSNL 60  
Db 67 HKVLTATAIQDATSQIKNTTPTVLTQDPQLGISFNSLSEITSTQTTILASTTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTPSKPTTKQRONKPNKPNNDFFHFEVFNVPSCISNNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTPSKPTTKQRONKPNKPNNDFFHFEVFNVPSCISNNPTCWAIC 186  
QY 121 KRIPNKPFGKTKTKPTKPTFKTKKDLKPQTKPKKEVPTTKTEBPTINTTKNTITTT 180  
Db 187 KRIPNKPFGKTKTKPTKPTFKTKKDLKPQTKPKKEVPTTKTEBPTINTTKNTITTT 246  
QY 181 LLTNNTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
Db 247 LLTNNTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298

## RESULT 5

US-08-467-969A-8  
; Sequence 8, Application US/08467969A  
; Patent No. 6168786  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasyszyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,969A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-475 MIS:bh  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-467-969A-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTATAIQDATSQIKNTTPTVLTQDPQLGISFNSLSEITSTQTTILASTTTPGVKSNL 60

Db 67 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFNVPCSCSNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFNVPCSCSNPTCWAIC 186  
QY 121 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 180  
Db 187 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 246  
QY 181 LTNNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSSPNTTRQ 232  
Db 247 LTNNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSSPNTTRQ 298

## RESULT 6

US-08-467-961A-8  
; Sequence 8, Application US/08467961A  
; Patent No. 6171783  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasysbyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6TH Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,961A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-476 MIS:bb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-467-961A-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 60  
|||||

Db 67 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFNVPCSCSNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFNVPCSCSNPTCWAIC 186  
QY 121 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 180  
Db 187 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 246  
QY 181 LTNNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSSPNTTRQ 232  
Db 247 LTNNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSSPNTTRQ 298

## RESULT 7

US-08-001-554A-8  
; Sequence 8, Application US/08001554A  
; Patent No. 6225031  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasysbyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/001,554A  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-286  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
US-08-001-554A-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 60  
Db 67 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFNVPCSCSNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVEFNVPCSCSNPTCWAIC 186  
QY 121 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 180  
Db 187 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 246

QY 181 LLTNTTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 |||||  
 Db 247 LLTNTTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 |||||

## RESULT 8

5194595-19  
 ; Patent No. 5194595  
 ; APPLICANT: WATHEN, MICHAEL W.  
 ; TITLE OF INVENTION: CHIMERIC GLYCOPROTEINS CONTAINING  
 ; IMMUNOGENIC SEGMENT OF THE GLYCOPROTEINS OF HUMAN RESPIRATORY  
 ; SYNCYTIAL VIRUS  
 ; NUMBER OF SEQUENCES: 19  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/543,780  
 ; FILING DATE: 31-OCT-1988  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 137,387  
 ; FILING DATE: 23-DEC-1987  
 ; SEQ ID NO:19:  
 ; LENGTH: 681  
 5194595-19

Query Match 73.3%; Score 901; DB 6; Length 681;  
 Best Local Similarity 91.0%; Pred. No. 1.2e-74;  
 Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 30 QLGISPSNLSEITSTQTTILASTTPGVKSLQPTTKTKNTTTTQTPGSKPTTKQKNKP 89  
 |||||  
 Db 490 QLGISPSNBEITSIQTTILASTTPGVKSLQPTTKTKNTTTTQTPGSKPTTKQKNKP 549  
 |||||  
 QY 90 ENKPNDFHFEVNFVPCISCSNNPTCWAICRIPNKKPKTKTKPTTKTKDOL 149  
 |||||  
 Db 550 PSKPNDFHFEVNFVPCISCSNNPTCWAICRIPNKKPKTKTKPTTKTKDOP 609  
 |||||  
 QY 150 KPQTTKEVPTTKPTKEPTINTTKNTTTTLLTNTTGNPKLTSQMETFHTSTSEGNLS 209  
 |||||  
 Db 610 KPQTTKEVPTTKPTKEPTINTTKNTTTTLLTNTTGNPKLTSQMETFHTSTSEGNLS 669  
 |||||

QY 210 PSQVSTTSE 218  
 |||||

Db 670 PSQVNISSQ 678  
 |||||

## RESULT 9

US-08-793-792-12  
 ; Sequence 12, Application US/08793792  
 ; Patent No. 6077511  
 ; GENERAL INFORMATION:

; APPLICANT:  
 ; TITLE OF INVENTION: Antigenic peptides derived from the  
 ; TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis  
 ; TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.  
 ; NUMBER OF SEQUENCES: 13  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/793,792  
 ; FILING DATE:

; CLASSIFICATION: 514  
 ; INFORMATION FOR SEQ ID NO: 12:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 37 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; HYPOTHETICAL: NO  
 US-08-793-792-12

Query Match 18.1%; Score 222; DB 3; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-14;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 NKPNNDFHFEVNFVPCISCSNNPTCWAICRIPNKK 127  
 |||||  
 Db 1 NKPNNDFHFEVNFVPCISCSNNPTCWAICRIPNKK 37  
 |||||

## RESULT 10

US-08-793-792-8  
 ; Sequence 8, Application US/08793792  
 ; Patent No. 6077511  
 ; GENERAL INFORMATION:

; APPLICANT:  
 ; TITLE OF INVENTION: Antigenic peptides derived from the  
 ; TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis  
 ; TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.  
 ; NUMBER OF SEQUENCES: 13  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/793,792  
 ; FILING DATE:

; CLASSIFICATION: 514  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 32 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; HYPOTHETICAL: NO  
 US-08-793-792-8

Query Match 15.7%; Score 193; DB 3; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-11;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 KPNDFHFEVNFVPCISCSNNPTCWAICR 123  
 |||||  
 Db 1 KPNDFHFEVNFVPCISCSNNPTCWAICR 32  
 |||||

## RESULT 11

US-08-928-361B-8  
 ; Sequence 8, Application US/08928361B  
 ; Patent No. 6071518  
 ; GENERAL INFORMATION:

; APPLICANT: Petersen, Carolyn  
 ; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
 ; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
 ; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
 ; TITLE OF INVENTION: SPECIES INFECTIONS  
 ; NUMBER OF SEQUENCES: 30  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
 ; STREET: 385 Sherman Avenue, Suite 6  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94306-1840

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/928,361B

```
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, HANA
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 216 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-928-361B-8

Query Match 15.3%; Score 188; DB 3; Length 216;
Best Local Similarity 26.4%; Pred. No. 8.3e-10;
Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;

QY 4 TLTAIIQDATSQIKNTTPTLTQDPQLGISFNSLSEITSTTILASTTPGVKSNLQPT 63
Db 12 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 71
QY 64 TVTKNTTTTTOTOPSKPTTKQKQKPNKPNNDPHEVFNVPSCSNNPTCWAICKRI 123
Db 72 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 111
QY 124 PNKKPGKKTTPKTKPTFKTKKOLKPKQTKPKKEVPTTKPTBPTINTKNTITLLT 183
Db 112 TTKKPTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 171
QY 184 NNTGNPKLTSQMETFHSSTSEGNLSQSVSTTSEHPSPQSSPNTT 230
Db 172 TTTTNTTTTTTATTTTKKPTTT-----TTTTTTTTPKPTTTTATTTT 214

RESULT 12
US-09-588-995A-8
; Sequence 8, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 08/700,651
; PRIOR FILING DATE: 1996-08-14
; PRIOR APPLICATION NUMBER: 08/415,751
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-8

QY 4 TLTAIIQDATSQIKNTTPTLTQDPQLGISFNSLSEITSTTILASTTPGVKSNLQPT 63
Db 12 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 71
QY 64 TVTKNTTTTTOTOPSKPTTKQKQKPNKPNNDPHEVFNVPSCSNNPTCWAICKRI 123
Db 72 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 111
QY 124 PNKKPGKKTTPKTKPTFKTKKOLKPKQTKPKKEVPTTKPTBPTINTKNTITLLT 183
Db 112 TTKKPTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 171
QY 184 NNTGNPKLTSQMETFHSSTSEGNLSQSVSTTSEHPSPQSSPNTT 230
Db 172 TTTTNTTTTTTATTTTKKPTTT-----TTTTTTTTPKPTTTTATTTT 214
```

```
Query Match 15.3%; Score 188; DB 4; Length 216;
Best Local Similarity 26.4%; Pred. No. 8.3e-10;
Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;

QY 4 TLTAIIQDATSQIKNTTPTLTQDPQLGISFNSLSEITSTTILASTTPGVKSNLQPT 63
Db 12 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 71
QY 64 TVTKNTTTTTOTOPSKPTTKQKQKPNKPNNDPHEVFNVPSCSNNPTCWAICKRI 123
Db 72 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 111
QY 124 PNKKPGKKTTPKTKPTFKTKKOLKPKQTKPKKEVPTTKPTBPTINTKNTITLLT 183
Db 112 TTKKPTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 171
QY 184 NNTGNPKLTSQMETFHSSTSEGNLSQSVSTTSEHPSPQSSPNTT 230
Db 172 TTTTNTTTTTTATTTTKKPTTT-----TTTTTTTTPKPTTTTATTTT 214

RESULT 13
US-08-928-361B-5
; Sequence 5, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, HANA
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1837 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-928-361B-5

Query Match 15.3%; Score 188; DB 3; Length 1837;
Best Local Similarity 26.4%; Pred. No. 1.4e-08;
Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;

QY 4 TLTAIIQDATSQIKNTTPTLTQDPQLGISFNSLSEITSTTILASTTPGVKSNLQPT 63
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:11:05 ; Search time 52.966 Seconds  
(without alignments)  
695.249 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVTLTAIQDTSQIKNT.....VSTTSEHPSPFPWTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03.\*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID          | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 1229  | 100.0       | 232    | 20 AAW96314 | Secreted G protein |
| 2          | 1229  | 100.0       | 298    | 14 AAR39286 | Respiratory syncyt |
| 3          | 1229  | 100.0       | 298    | 10 AAW96313 | Membrane bound G p |
| 4          | 1120  | 91.1        | 298    | 8 AAP70845  | Sequence of human  |
| 5          | 1120  | 91.1        | 298    | 13 AAR25302 | HSRV glycoprotein  |
| 6          | 1120  | 91.1        | 298    | 19 AAW47605 | HSRV glycoprotein  |
| 7          | 1120  | 91.1        | 298    | 23 AAU74676 | Respiratory syncyt |
| 8          | 901   | 73.3        | 681    | 10 AAP90441 | Chimeric human res |
| 9          | 580   | 47.2        | 361    | 24 ABP97862 | RSV G protein pept |

|    |       |      |     |             |                    |
|----|-------|------|-----|-------------|--------------------|
| 10 | 572.5 | 46.6 | 452 | 22 AAB67771 | Amino acid sequenc |
| 11 | 564.5 | 45.9 | 291 | 24 ABP97861 | RSV G protein pept |
| 12 | 564.5 | 45.9 | 548 | 24 ABP97863 | RSV G protein pept |
| 13 | 562   | 45.7 | 349 | 17 AAR95660 | Streptococcal prot |
| 14 | 562   | 45.7 | 349 | 22 AAB68028 | Amino acid sequenc |
| 15 | 558   | 45.4 | 101 | 15 AAR98253 | RSV subgroup A wil |
| 16 | 558   | 45.4 | 101 | 17 AAR95610 | RSV subgp. A prote |
| 17 | 558   | 45.4 | 101 | 17 AAR95616 | RSV sub-group A w1 |
| 18 | 558   | 45.4 | 101 | 17 AAR95616 | Respiratory Syncyt |
| 19 | 558   | 45.4 | 101 | 20 AAY44078 | RSV G protein anti |
| 20 | 558   | 45.4 | 101 | 21 AAB18805 | A G2Na peptide der |
| 21 | 558   | 45.4 | 101 | 22 AAG57741 | Amino acid sequenc |
| 22 | 558   | 45.4 | 101 | 22 AAB84123 | Amino acid sequenc |
| 23 | 558   | 45.4 | 101 | 22 AAB68016 | Amino acid sequenc |
| 24 | 558   | 45.4 | 101 | 22 AAB67775 | Human G protein, G |
| 25 | 552.5 | 45.0 | 356 | 17 AAR95661 | Immunogenic carrie |
| 26 | 538   | 43.8 | 101 | 16 AAR88255 | RSV subgroup A mod |
| 27 | 538   | 43.8 | 101 | 17 AAR95612 | RSV subgp. A prote |
| 28 | 538   | 43.8 | 101 | 17 AAR95618 | Respiratory Syncyt |
| 29 | 538   | 43.8 | 101 | 17 AAR97052 | RSV G protein anti |
| 30 | 538   | 43.8 | 101 | 20 AAY44080 | Peptide which indu |
| 31 | 538   | 43.8 | 101 | 22 AAB97311 | Amino acid sequenc |
| 32 | 538   | 43.8 | 101 | 22 AAB84125 | Human G protein, G |
| 33 | 538   | 43.8 | 101 | 23 AAO22581 | Peptide which indu |
| 34 | 538   | 43.8 | 101 | 23 AAO22581 | Human G protein, G |
| 35 | 538   | 43.8 | 101 | 23 AAO22586 | Human G protein, G |
| 36 | 534   | 43.4 | 101 | 20 AAW97310 | Peptide which indu |
| 37 | 534   | 43.4 | 101 | 23 AAO22585 | Human G protein, G |
| 38 | 514   | 41.8 | 101 | 20 AAW97312 | Peptide which indu |
| 39 | 514   | 41.8 | 101 | 23 AAO22587 | Human G protein, G |
| 40 | 506   | 41.2 | 101 | 17 AAR95614 | Amino acid sequenc |
| 41 | 506   | 41.2 | 101 | 17 AAR97063 | Respiratory Syncyt |
| 42 | 506   | 41.2 | 101 | 20 AAY44090 | RSV G protein anti |
| 43 | 506   | 41.2 | 101 | 22 AAB84135 | Amino acid sequenc |
| 44 | 506   | 41.2 | 101 | 22 AAB84135 | Human G protein, G |
| 45 | 506   | 41.2 | 101 | 23 AAO22584 |                    |

#### ALIGNMENTS

RESULT 1  
AAW96314  
ID AAW96314 standard; Protein; 232 AA.  
XX  
AC AAW96314;  
XX  
DT 28-JUN-1999 (first entry)  
XX  
DE Secreted G protein of respiratory syncytial virus.  
XX  
KW G protein; respiratory syncytial virus; RSV; recombinant vector;  
KW vaccine; immune response; immunogenicity; tPA; antibody;  
KW tissue plasminogen activator.  
XX  
OS Respiratory syncytial virus.  
XX  
PN WO9904010-A1.  
XX  
PD 28-JAN-1999.  
XX  
PF 16-JUL-1998; 98WO-CA00697.  
XX  
PR 18-JUL-1997; 97US-0896442.  
XX  
PA (CONN-) CONNAUGHT LAB LTD.  
XX  
PI Klein ME, Li X, Sambhara S;  
XX  
DR WPI; 1999-132254/11.  
XX  
DR N-PSDB; AAX08422.  
XX  
PT Immunogenic composition for generating antibodies against

PT respiratory syncytial virus - comprises non-replicating vector  
 PT containing the protein G sequence, useful in protective vaccines and  
 PT to raise antibodies for diagnosis  
 XX  
 PS Claim 9; Fig 3; 67pp; English.  
 XX  
 CC The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods. This truncated G  
 CC protein is secreted since it lacks a transmembrane domain.  
 XX  
 SQ Sequence 232 AA;  
 Query Match 100.0%; Score 1229; DB 20; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 3 6e-89;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HKVTLTAIIQDATSOIKNTTPTLYLTQDPQLGISFNSLSEITSTTTILASTTPGVKSNL 60  
 DB 1 HKVTLTAIIQDATSOIKNTTPTLYLTQDPQLGISFNSLSEITSTTTILASTTPGVKSNL 60  
 QY 61 OPTTVTKNTTTTQTPSKPTTKQKQNKPNKPNNDHFHEVFVPCISCSNNPTCWAIC 120  
 DB 61 OPTTVTKNTTTTQTPSKPTTKQKQNKPNKPNNDHFHEVFVPCISCSNNPTCWAIC 120  
 QY 121 KRIPNKKPKGKTTTKTKPTFKTKKOLKPKQTKKPKVEPTTKPTKEPTINTTKNITTT 180  
 DB 121 KRIPNKKPKGKTTTKTKPTFKTKKOLKPKQTKKPKVEPTTKPTKEPTINTTKNITTT 180  
 QY 181 LITNNTTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 DB 181 LITNNTTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 RESULT 2  
 AAR39286  
 ID AAR39286 standard; Protein; 298 AA.  
 XX  
 AC AAR39286;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 13-JAN-1994 (first entry)  
 XX  
 DE Respiratory syncytial virus (RSV) G protein.  
 XX  
 KW PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.  
 XX  
 OS Respiratory syncytial virus.  
 XX  
 PN WO9314207-A1.  
 XX  
 PD 22-JUL-1993.  
 XX  
 PF 05-JAN-1993; 93WO-CA00001.  
 XX  
 PR 06-JAN-1992; 92GB-0000117.  
 XX  
 PA (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PI Ewasysyn ME, Klein MH;  
 XX  
 DR WPI; 1993-243222/30.  
 DR N-PSDB; AAQ45686.

XX Multimeric hybrid genes and their chimeric proteins - are  
 PT vaccines against multiple pathogenic infections e.g.  
 PT para-influenza virus and respiratory syncytial virus  
 XX  
 PS Claim 11; Figure 7A-7D; 80pp; English.  
 XX  
 CC A novel multimeric hybrid gene is used as a vaccine. The gene  
 CC consists of two gene sequences which are linked and encode antigenic  
 CC regions, these two sequences being derived from two different  
 CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus  
 CC (RSV)). The gene sequences that are particularly used are those  
 CC which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
 CC G proteins (AAQ45685, AAQ45686).  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 298 AA;  
 Query Match 100.0%; Score 1229; DB 14; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 4 7e-89;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HKVTLTAIIQDATSOIKNTTPTLYLTQDPQLGISFNSLSEITSTTTILASTTPGVKSNL 60  
 DB 67 HKVTLTAIIQDATSOIKNTTPTLYLTQDPQLGISFNSLSEITSTTTILASTTPGVKSNL 126  
 QY 61 OPTTVTKNTTTTQTPSKPTTKQKQNKPNKPNNDHFHEVFVPCISCSNNPTCWAIC 120  
 DB 127 OPTTVTKNTTTTQTPSKPTTKQKQNKPNKPNNDHFHEVFVPCISCSNNPTCWAIC 186  
 QY 121 KRIPNKKPKGKTTTKTKPTFKTKKOLKPKQTKKPKVEPTTKPTKEPTINTTKNITTT 180  
 DB 187 KRIPNKKPKGKTTTKTKPTFKTKKOLKPKQTKKPKVEPTTKPTKEPTINTTKNITTT 246  
 QY 181 LITNNTTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 DB 247 LITNNTTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 RESULT 3  
 AAW96313  
 ID AAW96313 standard; Protein; 298 AA.  
 XX  
 AC AAW96313;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Membrane bound G protein of respiratory syncytial virus.  
 XX  
 KW G protein; respiratory syncytial virus; RSV; recombinant vector;  
 KW vaccine; immune response; immunogenicity; tPA; antibody;  
 KW tissue plasminogen activator.  
 XX  
 OS Respiratory syncytial virus.  
 XX  
 PN WO9904010-A1.  
 XX  
 PD 28-JAN-1999.  
 XX  
 PF 16-JUL-1998; 98WO-CA00697.  
 XX  
 PR 18-JUL-1997; 97US-0896442.  
 XX  
 PA (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PI Klein MH, Li X, Sambhara S;  
 XX  
 DR WPI; 1999-132254/11.  
 DR N-PSDB; AAX08421.  
 XX  
 PT Immunogenic composition for generating antibodies against  
 PT respiratory syncytial virus - comprises non-replicating vector  
 PT containing the protein G sequence, useful in protective vaccines and

PT to raise antibodies for diagnosis  
 XX  
 PS Claim 4; Fig 2; 67pp; English.  
 XX  
 CC The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods.  
 XX  
 SQ Sequence 298 AA;  
 Query Match 100.0%; Score 1229; DB 20; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-89;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HKVLTLLIADATSOIKNTTPTLTQDPQLGISFNSLSEITSTTTILASTTPGVKSNL 60  
 DB 67 HKVLTLLIADATSOIKNTTPTLTQDPQLGISFNSLSEITSTTTILASTTPGVKSNL 126  
 QY 61 OPTVTKNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 127 OPTVTKNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
 QY 121 KRIPNKKPGKKTTPKTKPKTKKDLKPQTKPKKEVPTTKTEPTINTKNTITTT 180  
 DB 187 KRIPNKKPGKKTTPKTKPKTKKDLKPQTKPKKEVPTTKTEPTINTKNTITTT 246  
 QY 181 LLTNNITGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSPSPPTNTTQ 232  
 DB 247 LLTNNITGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSPSPPTNTTQ 298  
 RESULT 4  
 AAP70845  
 ID AAP70845 standard; protein; 298 AA.  
 XX  
 AC AAP70845;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 05-APR-1991 (first entry)  
 XX  
 DE Sequence of human respiratory syncytial virus (HRSV) A2 strain  
 DE G protein.  
 XX  
 KW Vaccine.  
 XX  
 OS Human respiratory syncytial virus (HRSV).  
 XX  
 PN WO8704185-A.  
 XX  
 PD 16-JUL-1987.  
 XX  
 PF 23-DEC-1986; 86WO-US02756.  
 XX  
 PR 14-JAN-1986; 86US-0818740.  
 XX  
 PA (UYN-) UNIV NORTH CAROLINA.  
 PA (WERTZ) WERTZ G W.  
 XX  
 DR WP1; 1987-206300/29.  
 DR N-PSDB; AAN70784.  
 XX  
 PT Vaccines for human respiratory virus - comprising proteins or  
 PT fragment encoded by a DNA sequence coding for human respiratory

PT syncytial virus proteins.  
 XX  
 PS Disclosure; Chart 13; 57pp; English.  
 XX  
 CC A novel plasmid which comprises a DNA sequence encoding this  
 CC protein, and the protein itself, are claimed, for use as HRSV  
 CC vaccines. The vaccine can be administered to pregnant women or to  
 CC women of child bearing age to stimulate maternal antibodies.  
 CC Infants can also be vaccinated at 2-3 months of age.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 298 AA;  
 Query Match 91.1%; Score 1120; DB 8; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.8e-80;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 HKVLTLLIADATSOIKNTTPTLTQDPQLGISFNSLSEITSTTTILASTTPGVKSNL 60  
 DB 67 HKVLTLLIADATSOIKNTTPTLTQDPQLGISFNSLSEITSTTTILASTTPGVKSNL 126  
 QY 61 OPTVTKNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 127 OPTVTKNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
 QY 121 KRIPNKKPGKKTTPKTKPKTKKDLKPQTKPKKEVPTTKTEPTINTKNTITTT 180  
 DB 187 KRIPNKKPGKKTTPKTKPKTKKDLKPQTKPKKEVPTTKTEPTINTKNTITTT 246  
 QY 181 LLTNNITGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSPSPPTNTTQ 232  
 DB 247 LLTNNITGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSPSPPTNTTQ 298  
 RESULT 5  
 AAR25302  
 ID AAR25302 standard; Protein; 298 AA.  
 XX  
 AC AAR25302;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 03-MAR-1993 (first entry)  
 XX  
 DE HRSV glycoprotein G (gpG).  
 XX  
 KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;  
 KW major capsid protein; N.  
 XX  
 OS Human respiratory syncytial virus strain A2.  
 XX  
 PN US5149650-A.  
 XX  
 PD 22-SEP-1992.  
 XX  
 PF 13-JUL-1988; 88US-0218737.  
 XX  
 PR 14-JAN-1986; 86US-0818740.  
 PR 13-JUL-1988; 88US-0218737.  
 XX  
 PA (UYN-) UNIV NORTH CAROLINA.  
 XX  
 PI Collins PL, Wertz GW;  
 XX  
 DR WP1; 1992-340247/41.  
 DR N-PSDB; AAQ29623.  
 XX  
 PT Vaccines for human respiratory virus - include structural genes  
 PT coding for native structural viral proteins and immunogenic  
 PT fragments  
 XX  
 PS Disclosure; Page 18; 21pp; English.  
 XX  
 CC The sequences of mRNA encoding HRSV structural proteins are given in

CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid  
 CC protein N. The sequences and encoded proteins are useful for  
 CC preparing vaccines against HRSV. The vaccines can be used to confer  
 CC immunity against respiratory tract infections on human subjects.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX Sequence 298 AA;  
 SQ

Query Match 91.1%; Score 1120; DB 13; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.8e-80;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 HKVLTLLAIQDATSQIKNTTPTLYLTQDPQLGIFSNLSSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVPTTLLAIQDATSQIKNTTPTLYLTQDPQLGIFSNLSSEITSTQTTILASTTPGVKSTL 126

QY 61 QPTTVTKNTTTTQTPSKPTTKQKONKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 120  
 DB 127 QSTTVTKNTTTTQTPSKPTTKQKONKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 186

QY 121 KRIPNKKPGKTTTKPTKPTTKKDKLPQTTKPKKEVPTTKPTTEPTINTKTNIITT 180  
 DB 187 KRIPNKKPGKTTTKPTKPTTKKDKLPQTTKPKKEVPTTKPTTEPTINTKTNIITT 246

QY 181 LLTNNNTGNPKLTSMETFHSTSEGNLSPSQSVSTTSEHPSPSPSPNTTRQ 232  
 DB 247 LLTNTGPNELTSMETFHSTSEGNLSPSQSVSTTSEHPSPSPSPNTTRQ 298

RESULT 6  
 AAU47605  
 ID AAU47605 standard; Protein; 298 AA.  
 XX  
 AC AAU47605;  
 XX  
 DT 11-JUN-1998 (first entry)  
 DE HRSV glycoprotein G.  
 XX  
 KW HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine.  
 XX Human respiratory syncytial virus.  
 OS  
 PN US5716823-A.  
 XX  
 PD 10-FEB-1998.  
 XX  
 PF 12-MAY-1997; 97US-0854783.  
 XX  
 PR 13-JUL-1988; 88US-0218737.  
 PR 14-JAN-1986; 86US-0818740.  
 PR 23-DEC-1986; 86WO-US02756.  
 PR 11-JUN-1992; 92US-0897171.  
 PR 12-MAY-1997; 97US-0854783.  
 XX  
 PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Collins PL, Wertz GW;  
 XX  
 DR WPI; 1998-144802/13.  
 DR N-PSDB; AAV18736.  
 XX  
 PT Production of human respiratory syncytial virus glyco-protein F or G  
 PT - by culturing eukaryotic host cells transfected with corresponding  
 PT DNA  
 XX  
 PS Example 1; Columns 27-28; 17pp; English.  
 XX  
 CC The present sequence was used in the development of a novel method  
 CC for the production of human respiratory syncytial virus (HRSV)  
 CC glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises  
 CC culturing eukaryotic host cells transfected with an isolated DNA  
 CC sequence encoding HRSV gpF or gpG. The gp can be used to prepare

CC vaccines against HRSV.  
 XX Sequence 298 AA;  
 SQ

Query Match 91.1%; Score 1120; DB 19; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.8e-80;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 HKVLTLLAIQDATSQIKNTTPTLYLTQDPQLGIFSNLSSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVPTTLLAIQDATSQIKNTTPTLYLTQDPQLGIFSNLSSEITSTQTTILASTTPGVKSTL 126

QY 61 QPTTVTKNTTTTQTPSKPTTKQKONKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 120  
 DB 127 QSTTVTKNTTTTQTPSKPTTKQKONKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 186

QY 121 KRIPNKKPGKTTTKPTKPTTKKDKLPQTTKPKKEVPTTKPTTEPTINTKTNIITT 180  
 DB 187 KRIPNKKPGKTTTKPTKPTTKKDKLPQTTKPKKEVPTTKPTTEPTINTKTNIITT 246

QY 181 LLTNNNTGNPKLTSMETFHSTSEGNLSPSQSVSTTSEHPSPSPSPNTTRQ 232  
 DB 247 LLTNTGPNELTSMETFHSTSEGNLSPSQSVSTTSEHPSPSPSPNTTRQ 298

RESULT 7  
 AAU74676  
 ID AAU74676 standard; Protein; 298 AA.  
 XX  
 AC AAU74676;  
 XX  
 DT 09-APR-2002 (first entry)  
 DE Respiratory syncytial virus G protein.  
 XX  
 KW RSV; G protein; heavily glycosylated protein; antianaemic; antiviral;  
 KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;  
 KW cellular chemotherapeutic compound; humoral response;  
 KW cellular immune response; hPIV; paediatric respiratory disease;  
 KW globulin gene transfer; sickle cell disease; beta-thalassaemia;  
 KW human immunodeficiency virus infection; HIV.  
 XX Human respiratory syncytial virus.  
 OS  
 PN WO200192548-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 22-MAY-2001; 2001WO-US16610.  
 XX  
 PR 01-JUN-2000; 2000US-208701P.  
 XX  
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Portner A, Takimoto T;  
 XX  
 DR WPI; 2002-130534/17.  
 DR N-ESDB; AAS21045.  
 XX  
 PT Recombinant Sendai virus useful in vaccines to protect infection by  
 PT paramyxoviruses, comprises exogenous nucleic acid encoding  
 PT paramyxovirus protein or its antigenic fragment -  
 XX  
 PS Disclosure; Page 46; 57pp; English.  
 XX  
 CC The invention relates to a recombinant Sendai virus comprising an  
 CC exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its  
 CC antigenic fragment. The virus may be administered in combination  
 CC with an antiviral chemotherapeutic compound. Two or more viruses  
 CC expressing different PMV proteins may be co-administered. Compositions  
 CC comprising the virus are useful for eliciting a humoral and/or  
 CC cellular immune response to a PMV in a mammal, particularly a human.  
 CC Further a recombinant Sendai virus comprising an exogenous nucleic acid



CC anatoxin derived peptides, when modified to lack at least one cysteine  
 CC residue, are useful as carrier peptides. Deletion of Cys residues in  
 CC anatoxin peptides reduces formation of unwanted disulfide bridges. The  
 CC peptides are used as a carrier for vaccines, particularly those for  
 CC prevention or treatment of viral, bacterial, parasitic or fungal  
 CC infections, or cancers and to generate, or increase, an immune response  
 CC against infectious agents or tumour cells.

XX SQ Sequence 361 AA;  
 Query Match 47.2%; Score 580; DB 24; Length 361;  
 Best Local Similarity 70.8%; Pred. No. 8.3e-38;  
 Matches 114; Conservative 5; Mismatches 20; Indels 22; Gaps 3;

QY 64 TVTKNTNTTQTPSKPTTKQKONKPNKPNDFHFEVFNVPSCSNNPTCWAICKRI 123  
 Db 4 TVTKNTNTTQTPSKPTTKQKONKPNKPNDFHFEVFNVPSCSNNPTCWAICKRI 63  
 QY 124 PNKKPGKTTTKPKPTTKTKKDLKPKPTTKPKVPTTKPTEPTIN-----TTKT 175  
 Db 64 PNKKPGKTTTKPKPTTKTKKDLKPKPTTKPKVPTTKPTEPTIN-----TTKT 175  
 QY 176 NITT-----TLNNTTGNPKLTSQMETPHSTSE 205  
 Db 121 KIESLKEHGPINKMSESPNKTVSEKAKQYLEEFHQTALE 161

RESULT 10  
 AAB67771  
 ID AAB67771 standard; Protein; 452 AA.  
 AC AAB67771;  
 XX  
 DT 11-JUN-2001 (first entry)  
 DE Amino acid sequence of a fusion protein of P40 and RSV antigen.  
 KW Outer membrane protein A; OmpA; P40; enterobacteria; nasal composition;  
 KW RSV; RSV infection; lung; respiratory tract; vaccine.  
 XX Synthetic.  
 OS Klebsiella pneumoniae.  
 OS Respiratory syncytial virus.  
 XX WO200121203-A1.  
 XX 29-MAR-2001.  
 XX 22-SEP-2000; 2000WO-FR02626.  
 XX 23-SEP-1999; 99FR-0011888.  
 PA (FABR ) FABRE MEDICAMENT SA PIERRE.  
 PI Corvaiea N, Goestch L;  
 XX WPI; 2001-257929/26.  
 DR N-PSDB; AAF80153.  
 XX Vaccine against respiratory syncytial virus, comprises enterobacterial  
 PT outer membrane protein and viral immunogen, provides protective  
 PT response throughout the respiratory tract -  
 XX  
 PS Example 2; Page 31-32; 39pp; French.  
 XX The present sequence represents a fusion protein comprising a Klebsiella  
 CC pneumoniae outer membrane protein A (OmpA) designated P40 and a  
 CC respiratory syncytial virus (RSV) antigen. Enterobacterium OmpA proteins,  
 CC associated with an immunogenic peptide from RSV are used to prepare  
 CC a nasal composition that induces a protective response, against RSV  
 CC infection in the upper and lower (lung) respiratory tract. OmpA  
 CC potentiates the immune response to some immunogenic peptides,  
 CC eliminating the need for adjuvants. The method is useful for producing

CC vaccines for prevention or treatment of RSV infections.  
 XX SQ Sequence 452 AA;  
 Query Match 46.6%; Score 572.5; DB 22; Length 452;  
 Best Local Similarity 76.1%; Pred. No. 4.2e-37;  
 Matches 108; Conservative 5; Mismatches 18; Indels 11; Gaps 2;

QY 25 LTQDPQLGISFSLSEITSTQTTILASTTPGKSNLQPTTVTKTKNTTTTQTPSKPTTKQ 84  
 Db 322 LAPDRRVELEVKGKVVVQ-----PQPG-----DPMVTKTKNTTTTQTPSKPTTKQ 370  
 QY 85 RQNKPNKPNDFHFEVFNVPSCSNNPTCWAICKRIPNKKPGKTTTKPKPTFKT 144  
 Db 371 RQNKPNKPNDFHFEVFNVPSCSNNPTCWAICKRIPNKKPGKTTTKPKPTFKT 430  
 QY 145 TKKDLKPKPTTKPKVPTTKPTE 166  
 Db 431 TKKDLKPKPTTKPKVPTTKPVD 452

RESULT 11  
 ABP97861  
 ID ABP97861 standard; protein; 291 AA.  
 XX ABP97861;  
 XX  
 DT 03-JUN-2003 (first entry)  
 DE RSV G protein peptide G2Na and diphtheria anatoxin peptide DTa.  
 KW Diphtheria anatoxin; immune response; antigen; carrier; G protein;  
 KW hypersensitivity response; vaccine; infection; RSV; cancer.  
 XX Synthetic.  
 OS Respiratory syncytial virus.  
 OS Corynebacterium diphtheriae.  
 XX FR2827606-A1.  
 XX 24-JAN-2003.  
 XX 20-JUL-2001; 2001FR-0009733.  
 XX 20-JUL-2001; 2001FR-0009733.  
 PA (FABR ) FABRE MEDICAMENT SA PIERRE.  
 PI Corvaiea N, Nguyen NT, Beck A;  
 XX WPI; 2003-241969/25.  
 XX New peptide derived from diphtheria anatoxin, useful as carrier in  
 PT vaccines, lacks at least one Cys residue, also related nucleic acids -  
 PT Disclosure; Page 26-27; 42pp; French.  
 PS The present sequence represents a fusion protein comprising a peptide  
 CC derived from the G protein of Respiratory syncytial virus (RSV) linked  
 CC to a peptide derived from a diphtheria anatoxin. The diphtheria anatoxin  
 CC peptide generates an immune response against any antigen coupled to it,  
 CC and has a negative, immediate hypersensitivity response. Diphtheria  
 CC anatoxin derived peptides, when modified to lack at least one cysteine  
 CC residue, are useful as carrier peptides. Deletion of Cys residues in  
 CC anatoxin peptides reduces formation of unwanted disulfide bridges. The  
 CC peptides are used as a carrier for vaccines, particularly those for  
 CC prevention or treatment of viral, bacterial, parasitic or fungal  
 CC infections, or cancers and to generate, or increase, an immune response  
 CC against infectious agents or tumour cells.

XX SQ Sequence 291 AA;  
 Query Match 45.9%; Score 564.5; DB 24; Length 291;





Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;

QY 37 NLSEITSQTTT--ILASTT-----PGVKSNLQPTTVTKTKNTTTTQTPSKPTTKQKQK 89  
 Db 216 NGKTLGKETTTEAADAATARSFNFPILNSM---TVTKTKNTTTTQTPSKPTTKQKQK 272  
 QY 90 PNKPNDHFHEVFNFPVPCSCSNNTPTCWAICKRIPNKKPGKTKTKTKPTTKTKKDL 149  
 Db 273 PNKPNDHFHEVFNFPVPCSCSNNTPTCWAICKRIPNKKPGKTKTKTKPTTKTKKDL 149  
 QY 150 KPQTTKPKVEPTTKPTE 166  
 Db 333 KPQTTKPKVEPTTKPVD 349

RESULT 14  
 AAB68028  
 ID AAB68028 standard; Protein; 349 AA.  
 XX  
 AC AAB68028;  
 XX  
 DT 29-JUN-2001 (first entry)  
 XX  
 DE Amino acid sequence of fusion protein comprising 2 G protein fragments.  
 XX  
 KW Aliphatic ammonium salt; immunogen; antigen; syncytial virus infection;  
 KW G protein; fusion protein.  
 XX  
 OS Synthetic.  
 OS Streptococcus sp.  
 OS Respiratory syncytial virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 120..230  
 FT /note= "G protein fragment of VRS"  
 XX  
 XX FR2798292-AL.  
 XX  
 PD 16-MAR-2001.  
 XX  
 XX 09-SEP-1999; 99FR-0011284.  
 XX  
 XX 09-SEP-1999; 99FR-0011284.  
 XX  
 XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX  
 PI Beck A, Klinguer C, Nguyen TN;  
 XX  
 XX WPI; 2001-267782/28.  
 XX  
 DR N-PSDB; AAF84711.  
 XX  
 XX Use of quaternary aliphatic ammonium salt and immunogen or antigen to  
 PT combat respiratory syncytial virus infections -  
 XX  
 XX Claim 16; Page 22-23; 35pp; French.  
 XX  
 CC The specification describes a method which uses quaternary aliphatic  
 CC ammonium salts together with an immunogen or antigen to treat syncytial  
 CC virus infections. The combination of the salt with the antigen or  
 CC immunogen improves immunogenicity and equilibrates the Th1/Th2 immune  
 CC response. The method is used for the treatment of respiratory syncytial  
 CC virus infections. The present sequence represents a fusion protein,  
 CC comprising an albumin binding domain of the G protein of Streptococcus  
 CC sp. fused to a G protein fragment of respiratory syncytial virus (VRS,  
 CC long version). The fusion protein is used as an antigen in the method  
 CC of the invention.  
 XX  
 XX Sequence 349 AA;  
 SQ

Query Match 45.7%; Score 562; DB 22; Length 349;  
 Best Local Similarity 78.1%; Pred. No. 2.1e-36;  
 Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;

QY 37 NLSEITSQTTT--ILASTT-----PGVKSNLQPTTVTKTKNTTTTQTPSKPTTKQKQK 89  
 Db 216 NGKTLGKETTTEAADAATARSFNFPILNSM---TVTKTKNTTTTQTPSKPTTKQKQK 272  
 QY 90 PNKPNDHFHEVFNFPVPCSCSNNTPTCWAICKRIPNKKPGKTKTKTKPTTKTKKDL 149  
 Db 273 PNKPNDHFHEVFNFPVPCSCSNNTPTCWAICKRIPNKKPGKTKTKTKPTTKTKKDL 149  
 QY 150 KPQTTKPKVEPTTKPTE 166  
 Db 333 KPQTTKPKVEPTTKPVD 349

RESULT 15  
 AAR8253  
 ID AAR8253 standard; peptide; 101 AA.  
 XX  
 AC AAR8253;  
 XX  
 DT 15-MAY-1996 (first entry)  
 XX  
 DE RSV subgroup A wild type residues 130-230 (G2A clone).  
 XX  
 KW Immunogenic peptide; protein G; respiratory syncytial virus; adjuvant;  
 KW carrier protein; membrane lipopolysaccharide; LPS; Klebsiella pneumoniae;  
 KW divalent cation; detergent; anion-exchange chromatography; infection.  
 XX  
 OS Respiratory syncytial virus.  
 XX  
 XX FR2718452-AL.  
 XX  
 PD 13-OCT-1995.  
 XX  
 PF 06-APR-1994; 94FR-0004009.  
 XX  
 PR 06-APR-1994; 94FR-0004009.  
 XX  
 XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX  
 PI Bausant T, Binz H, Thien NN, Trudel M;  
 XX  
 DR WPI; 1995-353189/46.  
 DR N-PSDB; AAT03486.  
 XX  
 XX New respiratory syncytial virus polypeptide(s) for vaccine prodn. -  
 PT esp. by conjugation with new Klebsiella pneumoniae p40 protein  
 XX  
 XX Disclosure; Page 23; 38pp; French.  
 XX  
 CC The sequence of the wild type amino acid residues 130-230 from the  
 CC protein G of the respiratory syncytial virus (RSV) subgroups A. The  
 CC sequence was used to synthesise the immunogenic peptides AAR8245-52  
 CC based on residues 174-187 of the subgroup A and B RSVs. The peptides are  
 CC pref. conjugated to a novel carrier protein (the p40 protein; see  
 CC AAR8257) derived from a membrane lipopolysaccharide (LPS) fraction from  
 CC Klebsiella pneumoniae. The LPS fraction was isolated by precipitating the  
 CC Klebsiella membrane LPSs with a divalent cation and detergent,  
 CC subjecting the recovered proteins to anion-exchange chromatography to  
 CC obtain an immunological adjuvant and linking the p40 protein to the above  
 CC peptides. The conjugates are useful in the treatment of RSV A or B  
 CC infections.  
 XX  
 XX Sequence 101 AA;  
 SQ

Query Match 45.4%; Score 558; DB 16; Length 101;  
 Best Local Similarity 99.0%; Pred. No. 1e-36;  
 Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 64 TVTKTKNTTTTQTPSKPTTKQKQKPNKPNDHFHEVFNFPVPCSCSNNTPTCWAICKRI 123  
 Db 1 TVTKTKNTTTTQTPSKPTTKQKQKPNKPNDHFHEVFNFPVPCSCSNNTPTCWAICKRI 60  
 QY 124 PNKKPGKTKTKTKTKTKTKDLKPQTTKPKVEPTTKP 164

Db 61 PNNKPGKTTTKPKRPTFKTKRKHKPOFTKPKKEVPTTKP 101

Search completed: October 29, 2003, 17:38:01  
Job time : 53.966 secs